DESIGN AND APPLICATIONS OF DNA-BASED DEVICES FOR SELF-ASSEMBLY, MOLECULAR CIRCUITS, AND SOFT MATERIALS

by
Joshua Fern

A dissertation submitted to Johns Hopkins University in conformity with
the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
May, 2017

Abstract

Biologically inspired synthetic materials have led to novel technologies due of their ability to sense, influence, or adapt to their environment. One way to build these materials and devices is to utilize the high sequence specificity and innate biocompatibility of DNA. While once considered as a material useful for only storing genetic information, DNA-based devices are now being realized as molecular tools in fields such as therapeutics, diagnostics, regenerative medicine, and soft robotics. In this dissertation, we investigate the use of DNA to build programmable tools to control self-assembly, implement molecular computation, and direct material change processes.

DNA origami nanostructures are useful tools for controlling the spatial patterns of proteins, nanoparticles, and fluorophores because they contain hundreds of independently functionalizable locations that can be engineered with nanoscale precision. However, the addressable surface area is currently limited by the size of single origami structures, and efficient, high-yield self-assembly of multiple origami into higher-order assemblies continues to be a challenge. To investigate the factors important for heterogeneous self-assembly of multiple origami, we experimentally measure the equilibrium distribution of four origami tiles in the monomer, intermediate, and final tetramer states as a function of temperature. We find that the thermodynamics of the self-assembly process is determined by the binding interface between origami. Simulations of the assembly kinetics suggest assembly occurs primarily via hierarchical pathways.

Next, we engineer a DNA-based timer circuit that can be used in computational devices for molecular release or material control. The circuit releases target DNA sequences into solution at a programmable time with a tunable, constant rate. Multiple timer circuits can operate simultaneously, each releasing their target sequences at independent rates and times.

We further develop the utility of the timer and similar DNA-based circuits as a means to control molecular events in biological environments, such as serum-supplemented cell media, where DNA-degrading nucleases can reduce the functional stability and lifetime of DNA-based devices. By implementing DNA circuit-protective design principles and by adding screening molecules to reduce nuclease activity, the functional lifetime of simple DNA circuits can be significantly increased. We develop a model by fitting parameters for reactions between nucleases and simple DNA circuits. Using the model, we can qualitatively predict the behavior of more complex circuits: multiple circuits in series and circuits containing competitive reactions.

Finally, we investigate how DNA-based circuits can be used to trigger the high-degree swelling response of DNA-crosslinked metamorphic hydrogels. By coupling signal amplification to the triggering process, we demonstrate modular control over the timescale and degree of swelling. Further, we show control over the identity of the trigger molecule using molecular translators and computational controllers capable of converting complex chemical inputs into mechanical actuation.

Thesis Advisor: Dr. Rebecca Schulman

Thesis Readers: Dr. Sharon Gerecht, Dr. Honggang Cui, Dr. Margaret Johnson, Dr. Hai-Quan Mao

Alternate Readers: Dr. Jeffrey Gray, Dr. Winston Timp

Knowledge will forever govern ignorance: And a people who mean to be their own Governors must arm themselves with the power which knowledge gives.

-James Madison, 1822

Acknowledgements

After visiting Johns Hopkins during visitation weekend, working with Professor Rebecca Schulman was not much of a consideration. DNA as a building material for nanostructures? Sounds weird. What stuck with me was her energy, how kind she was, and how excited her first Ph.D. student, Abdul, was to work with her and on his 'peculiar' research. After listening to and meeting with the various professors during the advisor selection process, I knew that I wanted to work with Prof. Schulman. She was exciting, the research was new and exciting, and she took the time and energy to focus on her students and their research. I have not been disappointed and am very grateful for the choice I made. Prof. Schulman has been motivational, understanding, patient, has challenged me intellectually, and has been an integral part of my Ph.D. career.

Thank you to the amazing lab mates I have had over the past years. Whether we're discussing politics, the next big stock to invest in, watching SpaceX land their rockets, or debating the finer details of the latest journal article or technology, it has always been a joyful and a grateful distraction from the frustrations of the lab. I truly believe our lab culture was a large part of my research progression and I hope that the excellent sense of unity, collaboration, friendliness and comradery we have established will continue in the future. Special thanks to Ankur Verma, Abdul Majeed Mohammed, John Zenk, Angelo Cangialosi, and Sam Schaffter for not only being good lab mates, but being great friends outside of the lab. Thank you also to the students, Jen Lu and Chad Fisher, that I was lucky enough to mentor. Both of you, and many other JHU undergraduates who passed through our lab, are brilliant individuals and will go far in life. I doubt I will work with such a special group again.

Next, I'd like to thank my friends for all of their constant support and for keeping me grounded in the real world. The friends I have gained here in Baltimore and those from

Wisconsin and elsewhere have been a constant source of fun and adventure. Special shout out to David Raciti, whose stories were always fun to listen to and were a constant reminder that things could always be worse.

Finally, thank you to my parents and the rest of my family for the constant love and support they have given me through my college years, and especially to my grandparents for always being "so proud of me." I will forever be grateful to my parents for giving me the genes, tools, and opportunities to pursue a Ph.D.

Table of Contents

List of Figures

List of Tables

CHAPTER 1: INTRODUCTION

The purpose of DNA was once considered to be solely for the code to life and the machinery of cells, and exclusively belonging to the realm of biologists. Thirty years ago, that purpose started to change when Nadrian Seeman saw DNA as more than a tool for biologists. He envisioned DNA as a building block for nanoscale structures and other materials.[1,2] Taking lessons from DNA's use in biology, Seeman and his coworkers took advantage of the specificity of Watson-Crick sequence complementarity and Holliday junctions as a possible structural motif to build a synthetic lattice junction, and the first DNA nanostructure (Figure 1.1).[3] From these first junctions, they engineered the first synthetic multi-component DNA nanostructure, a cube.[4]

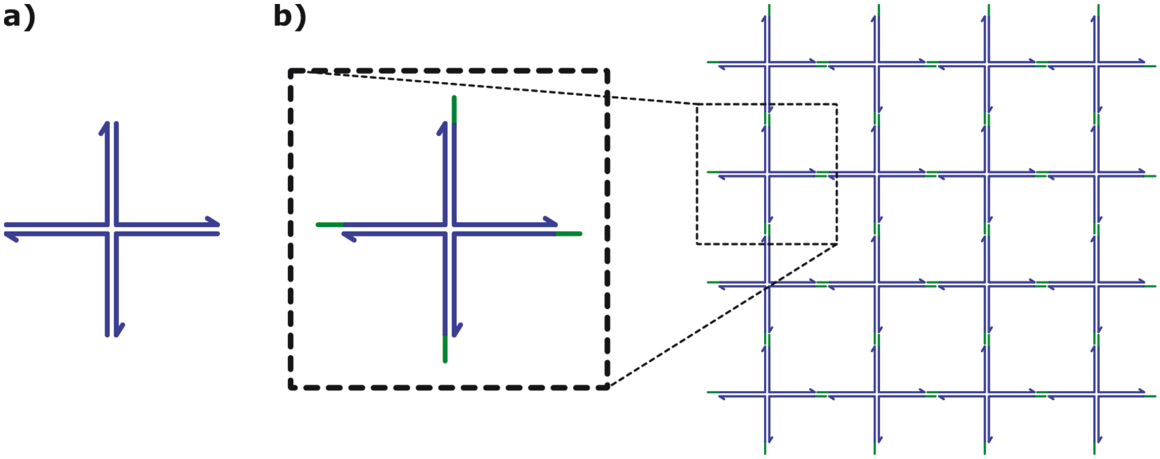


Figure 1.1: (a) Design for a DNA junction similar to the ones first constructed in 1983.[3] (b) Green "sticky-ends"[4] allow for the assembly of individual junctions into lattices or other multi-junction structures.[5]

Since Seeman's first demonstrations, the field of DNA nanotechnology has undergone incredible expansion and development into not only structural nanotechnology and self-assembly, but also the sub-fields of molecular computation, soft materials, and has, in a circular fashion, generated new tools to aid in studying biology or interacting with biological specimens. To aid the field's growth, the synthesis of DNA oligonucleotides and DNA plasmids has drastically decreased in cost over the past two decades.[6] This decrease in cost and the increased availability of modified- or chemically labeled-DNA fueled the engineering of computational and structural systems built with 100's of DNA strands, each with a unique, user-defined sequence.[7,8] The most notable nanostructures to arise from this technological growth are DNA origami.[7]

In 2006, Paul Rothemund published a landmark paper demonstrating the concept of DNA origami.[7] In DNA origami, short single-stranded DNA "staples," about 32 nucleotides in length, are designed to have sequence complementarity to regions of a long single-stranded DNA, or "scaffold" strand. The final shape of the origami is designed by raster-pattern of the scaffold strand, whose form is held together by the staple strands. When the staple and scaffold strands are annealed together in a salt-containing buffer from 90 to 20 °C, the strands will self-assemble, as predetermined by the sequences of the staple and scaffold strands, into the pre-specified shape (Figure 1.2). Since its conception, the DNA origami technique has been used to form a wide variety of 2D[7,9] and 3D[10–13] nanostructures. Similar techniques that rely on the self-assembly of only short single-stranded DNA oligonucleotides[8,14] or of DNA tile nanostructures[15,16] have also been developed.

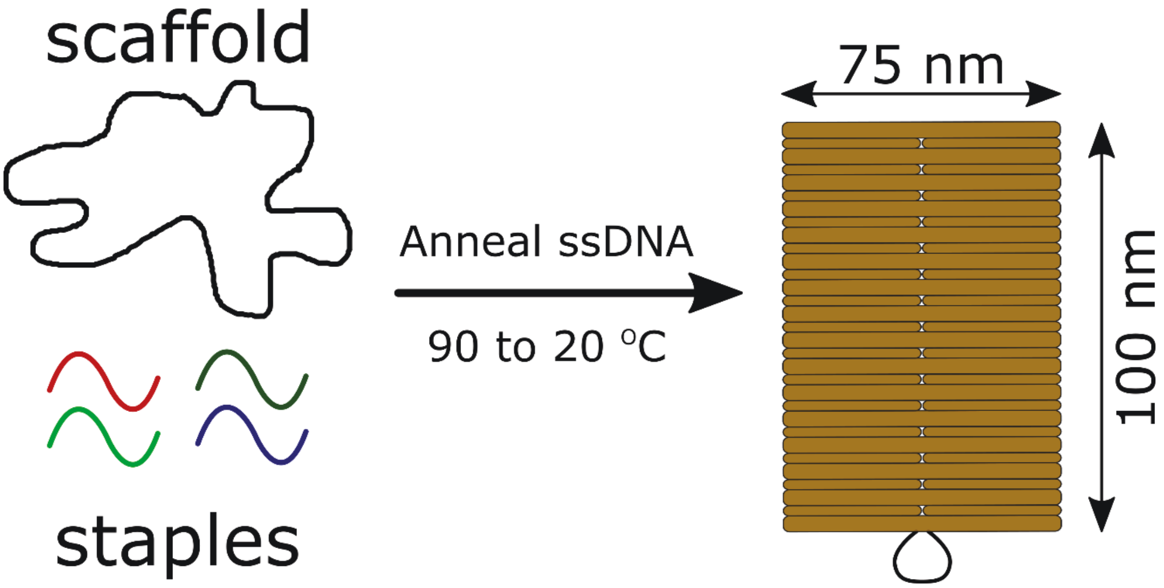


Figure 1.2: DNA origami are formed from the self-assembly of staple and scaffold strands. The single-stranded DNA are mixed in a salt-containing buffer and annealed, generally from 90 to 20 °C. A rectangle shape is shown as one possible origami shape that could be designed.

One of the major motivations behind DNA origami and associated nanostructures is the ability to engineer surfaces or objects where multiple locations on the structure are independently functionalizable due to the sequence specificity of DNA. In 2D DNA origami structures, the spatial resolution of the sequence-specific functionalizable locations is as low as 6 nm.[7] The functionalization of DNA nanostructures has led to the design of origami patterned with polymers,[17–19] fluorophores,[20,21] nanoparticles,[22] and proteins.[23–26] Functionalized nanostructures have been used as tools in cellular systems,[27–32] in colloidal self-assembly and plasmonics,[33–36] surface patterning,[37–39] and the formation of metallic nanoparticles or other inorganic materials.[40–42] However, the maximum surface area with origami folded from the original M13mp18 plasmid is currently only about 100 x 75 nm.[43,44] Though larger structures are possible by extending plasmid sequences *via* cloning,[45] surface area size scalability in this manner is ultimately limited by DNA synthesis cost and feasible maximum size of the plasmid.

1.1: Self-assembly of DNA nanostructures

To circumvent limitations on nanostructure size, DNA nanostructure components can be designed and programmed to assemble into multi-component structures, such that the resulting surface area or volume for functionalization is significantly greater than that of the original monomers. In nature, multi-component self-assembly of structures is critical for the functional operation and production of virus capsids,[46–48] cell machinery,[49–51] and biological tissue assembly.[52–54] In self-assembly, the monomer components are engineered or evolved to have domains that enable inter-component linking. When multiple components are mixed together, the components spontaneously bind together to assemble into multi-component structures. This principle has driven the engineering of synthetic, self-assembling structures built from peptides,[55,56] proteins,[57,58] and DNA.

The field of self-assembly is further divided into two sub categories: heterogeneous and homogeneous self-assembly. In homogeneous self-assembly, each monomer has the same structure and binding sites, such that when monomers are mixed together, they assemble into a 1, 2, or 3D lattice structure.[59–62] While homogeneous assembly requires the engineering of a single monomer component, the degree to which an area or volume is uniquely and independently patterned is limited because each monomer unit is the same. Thus homogeneous self-assembly is useful for building patterns or functions with repeating units in the defined assembly directions.[63,64]

In contrast to homogeneous self-assembly, heterogeneous self-assembly involves the assembly of few to many unique monomer components that may have unique or the same binding domains within each component. Heterogeneous self-assembly of finite multi-component assemblies requires binding interactions to be weak and reversible under assembly conditions in order to prevent kinetic trapping of components in intermediate states instead of the final structure.[65,66] Interactions must also be specific to prevent undesired assemblies from forming.[66–69] Heterogeneous self-assembly with DNA has been used to build arrays of repeating units[70–73] and structures that can conduct computation using templated assembly methods.[74–76] Historically, multi-component DNA origami structures have suffered from low yields due to inefficiencies in assembly and components becoming stuck in intermediate complexes.[77,78]

The binding domains used for the assembly of DNA nanostructures consists of either blunt-ends, sticky-ends, or a mixture thereof. A blunt-end occurs when two DNA strands of a double-stranded complex terminate without single-stranded DNA overhangs. If a single-stranded DNA overhang is present, and is complementary to a single-stranded sequence on a different component's binding site, then it is termed a "sticky-end."[4] Single-stranded overhangs could also be used to prevent blunt-end stacking between components by inhibiting $\pi - \pi$ stacking interactions between two DNA double-helices on two interacting components.[79] More recently, success in limiting non-specific interactions while maintaining a high-degree of specificity has been obtained by using combinations of blunt-end and sticky-end interactions to create relatively weak, but specific interactions.[80] These types of interfaces have led to the building of large crystal structures with a surface-bound size on the order of 10's of microns.[81]

One reason large scale multi-component DNA origami nanostructures have suffered from low yields is the lack of quantitative computational tools to drive reasonable component and interface design. Instead, the design of nanostructures for assembly purposes has primarily been experimentally driven, rather than by designs driven by computational predictions. Additionally, currently available software tools used to build the sequence level or architectural parts of the origami designs were not built with multi-component structures in mind.[10,82] To computationally design multi-component origami assembly, the simulation and design tools will need to incorporate the energetics of DNA origami and other nanostructure multi-component assembly[44,83] and the role local curvature plays in the global curvature of the super-structure.[11,60,81] More recent computational tools promise to reduce the limitations on computer-aided design and simulations,[84,85] and further experiments can be used to further drive the improvement of the available software.

1.2: DNA as a tool for molecular computing

In addition to the self-assembly of biomolecules, living systems are founded upon the concept of molecular computing where input sensing, the decision process, and the output action is conducted and orchestrated by biomolecules. For example, cells use chemical moieties on their surface to sense their surroundings. Once triggered, these sensors transduce the signal through multiple layers of internal machinery that results in decisions such as migration, cell death, proliferation, and the production of proteins or other chemical signals.[86] DNA computing is a subfield of DNA nanotechnology that focuses on applications of DNA in molecular computing, where information or signals are transduced, translated, stored, or calculated using DNA. The core of DNA computing is founded upon DNA strand-displacement reactions, where an invading single-stranded DNA species interacts with a double-stranded DNA complex and "displaces" a strand in the complex (Figure 1.3). DNA hybridization is primarily governed by the specificity of Watson-Crick base pairing such that combinatorial sequence design using the four bases is sufficient for large, multi-reaction strand-displacement systems. This has enabled the engineering of computational circuits such as a circuit that utilizes 130 unique DNA oligonucleotide strands (74 DNA species without inputs) to calculate the square-root value within a single reaction tube.[87] One of the major limitations of DNA computing is the relatively slow reaction rate constants of the reactions driving computational processes compared to silicon-based calculators. While a simple square root calculation takes fraction of a second for a hand-held calculator, the same operation requires hours with a DNA-based calculator.[87]

However, DNA-based computers, or more generally any implementation of a DNA strand-displacement reaction, can perform functions a hand-held calculator cannot. For example, DNA strand-displacement processes have become important cornerstones for the design of chemosensors,[88,89] directors of self-assembly,[90,91] drug therapy,[92] and diagnostics.[89,92,93] Additionally, DNA strand-displacement circuits are highly sequence specific,[94,95] scalable,[87] and tunable with controllable reaction rate constants ranging over 6 orders of magnitude.[96] Strand-displacement reactions are usually initiated *via* a single-stranded region termed a "toehold."[96] Once an invading strand binds a toehold (Figure 1.3a), branch migration, a series of random dissociation/association steps, can occur where the invading and incumbent strands compete for hybridization sites until one of the strands falls off or is displaced. Since invading strands have the additional thermodynamic stability *via* hybridization with the toehold, there is a higher propensity for incumbent strands to be fully displaced to complete the strand-displacement process.

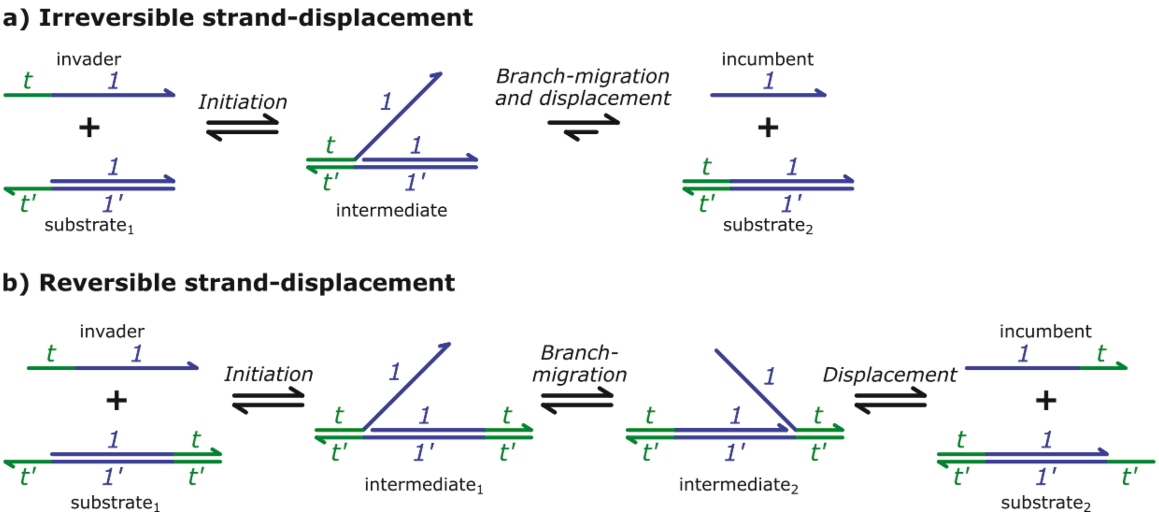


Figure 1.3: Schematic of irreversible (a) and reversible (b) strand-displacement reactions. The binding of the invader strand to the toehold (t') of the substrate initiates the strand-displacement process.

Reversible strand-displacement reactions correspond to reactions where both the invading and incumbent strand are complementary to a toehold on the same substrate (Figure 1.3b), and have been used to build a class of strand-displacement systems implementing "seesaw gates."[87,97] In addition to engineered toehold-mediated strand-displacement reactions, toehold-free (*i.e.*, 0 bp) reactions can occur and have been recognized as either a source of undesired "leak" reactions[96,98–100] or as method to release a strand at a constant, slow rate as shown in *Chapter 3*.[101] Zero base pair reactions are initiated by the random "breathing" or fraying of a double-stranded complex, predominantly at the termini, that temporarily creates a binding site for the invading strand.[96,98,100]

As DNA hybridization is nucleation-limited,[102] the kinetics of toehold-mediated strand-displacement reactions are primarily controlled by the number of bases in the toehold.[96] For a mass-action kinetics model of the reaction, the reaction rate constant increases by an order of magnitude for each base added to the toehold for toehold lengths between 0 and 6 nucleotides (*e.g.*, 0.5 to 5×10^5 1/M-sec).[96,98] A reaction using a 7 bp toehold has a rate constant of 1×10^6

1/M-sec, only 2-fold greater than that of a 6 bp toehold. Beyond 7 bp, the rate constant plateaus because the rate of dissociation upon initial toehold nucleation is essentially zero.[98] The kinetics of the strand-displacement process is dependent on the sequence of both the toehold and branch-migration domain,[96,102,103] and the choice of sequence can additionally influence the propensity for potential leak reactions in a circuit. The kinetics can also be tuned using base mismatches,[104,105] and base or backbone modifications.[106] By relying on sequence specificity and reaction rate tunability, strand-displacement circuits have been demonstrated to perform complex Boolean logic,[87,97,107,108] signal amplification,[109–111] and have been used to build reaction-diffusion mediated patterns within hydrogels.[112–115]

To aid in the building of these circuits, software packages can be used to calculate the salt concentration and temperature-dependent equilibrium structure of a collection of strands,[116] and software that simulates the strand-displacement kinetics of the proposed system to determine system behavior and the effect of potential leak reactions on system dynamics.[117–119] With increasing accuracy of DNA synthesis procedures and optimum sequence prediction capabilities of DNA strand-displacement modeling software, the propensity for leak reactions in experimental applications will decrease and their effect on the operability of strand-displacement systems will become better understood for the further improvement of reaction design.

1.3: DNA as a tool for the study of biology

One of the primary applications of DNA-based nanotechnologies lies within the study of biology. This has partially been led by the discovery of aptamers, DNA or RNA sequences that bind proteins, peptides, small molecules, or ions. Aptamers are designed using an *in vitro* evolution technique termed SELEX, or the systematic evolution of ligands by exponential enrichment,[120] and has driven the development of DNA-based therapeutics,[121] drug delivery devices,[122,123] and diagnostics[124] that are highly specific to a target molecule or cell species. The usefulness of aptamers as a treatment option is demonstrated by FDA approved aptamer-based therapies such as Macugen for the treatment of age-related macular degeneration in 2004.[125]

DNA can also be used as a tool to precisely spatially resolve cellular structures. The conjugation of DNA to specific biomolecules or antibodies enables the targeting of the desired protein or structure for analysis using fluorescence microscopy. Techniques such as DNA-PAINT and DNA-Exchange can be used with super-resolution microscopy analyses methods to discern multiple nanometer-scale cell structures on the single-cell and whole tissue level.[126] Larger nanostructures, such as DNA origami, have been used to not only label internal and external cell structures,[127,128] but to also direct cellular behavior and function.[28]

Inside cells, nuclease-mediated digestion of DNA, specifically by DNases, is an important mechanism for cell function.[129] Any DNA nanostructure or strand-displacement circuit added to this environment is susceptible to digestion conditions that could reduce the functionality of the nanostructure or circuit. Within the DNase class of nucleases, exonucleases are primarily responsible for removing bases from single-stranded DNA or from a blunt-end, and endonucleases are primarily responsible for cleaving polynucleotide chains internally. In synthetic biology, restriction enzymes are engineered or evolved endonucleases used to cleave at specific sequences. The most well-known human endonuclease, DNaseI, cleaves relatively non-specifically and is capable of digesting both single- and double-stranded DNA.[130] Because nucleases are present in serum often used as a cell culture supplement,[131] in blood,[132] and within cells themselves,[133] the functionality of DNA-based technologies is at risk of being disrupted or made completely inert. Many DNA and RNA modifications have been developed to combat nuclease-mediated degradation in biological samples, including the inverted dT,[134] phosphorylation, phosphorothioate backbones,[135] and 2' O-methyl bases.[136] However, most of these modifications are not scalable to use with DNA origami or other smaller DNA nanostructures due to the high cost of synthesis or the large, negative impact on the thermodynamics of DNA hybridization (*i.e.*, phosphorothioate bonds). DNA nanostructures themselves appear to have some resistance against digestion, potentially due to the packing of the DNA helices or steric effects that reduces nuclease affinity for the nanostructure.[137–139] Thus, efficient and affordable methods are still needed to protect DNA strand-displacement circuits from digestion.

In addition to interfacing with cells in cell media, DNA has also been used to direct cell behavior through the cell's interaction with a surface. Cells adhere to their surrounding environment through a multitude of mechanisms and proteins.[140] Since DNA is highly functionalizable, it can be bioconjugated to peptides or proteins to facilitate cell adhesion. For example, RGD functionalized DNA has been coated on surfaces to direct and investigate cell adhesion.[141,142] Additionally, surfaces micropatterned with DNA can serve as an array of attachment points to facilitate high-throughput screening of cell adhesion or behavior-influencing molecules.[143]

Another method to influence cell behavior through their adhesion complexes is by engineering hydrogels integrated with DNA.[144,145] Cell adhesion to these materials is facilitated by functionalizing the DNA with cell adhesion ligands, such as RGD. Using this method, the mechanical stiffness of the hydrogel can be dynamically controlled through time by the addition of DNA strands that soften or stiffen the gels by converting the crosslinker between single-stranded and double-stranded states.[146,147] This has been used to direct cell behavior[146,148] or direct the differentiation of stem cells into specific final cell types.[149] These techniques could lead to improved culture methods and diagnostic technologies.

1.4: DNA as a tool to control material properties

DNA has also been used to build unique soft material systems where DNA makes up either the entirety of the material, or is used as the crosslinker between non-nucleic acid polymer backbones. These materials have shown promise in applications such as shape memory,[150,151] drug delivery and therapeutics,[152] and chemical sensors.[153,154] As mentioned in Section 1.3, the mechanical properties (*e.g.*, stiffness) of the DNA-integrated materials can be dynamically tuned by adjusting crosslinker stiffness,[155] crosslinker density,[144] or by using DNA crosslinks whose secondary structures depend on ion concentration or pH.[156] Additional sensing of chemical information in these systems is realized using aptamers.[153,154] Thus far, dynamic changes to the DNA crosslinks primarily result in either low degrees of swelling or in irreversible dissociation of the crosslinks that limit potential applications in the field of soft-robotics and as actuatable materials.

This dissertation investigates applications implementing the most fundamental tool in DNA nanotechnology, DNA-DNA hybridization, and demonstrates how it can be used to engineer the controlling mechanisms behind self-assembly, the operation of molecular circuits, and the manipulation of polymeric soft materials. In *Chapter 2*, we use fluorescence colocalization microscopy to calculate the free energy change of the DNA hybridization-mediated self-assembly of four origami tiles into a heterogeneous tetrameric ring structure. The free energy changes are used to suggest the major influencers of self-assembly in this origami system and the most prominent pathways to assembly of the final tetramer product. *Chapter 3* transitions from large DNA origami structures to the smaller single-stranded and double-stranded nanostructures of molecular circuits in order to build a molecular timer that releases an output DNA molecule at independently tunable rates and delay times. Using simulations, we investigate the major modes of circuit non-ideality prior to and during output release. In *Chapter 4*, we build upon the simulations and DNA strand-displacement controllers built in *Chapter 3* to introduce and investigate possible DNA modifications and chemical supplements that enable reliable and robust functional operation of DNA-based circuits in serum-supplemented cell media – conditions under which circuit components can be digested by nucleases. In *Chapter 5*, we use our DNA-based circuits to control the activation state of DNA-crosslinked hydrogels capable of undergoing a large degree of swelling. We use modular DNA strand-displacement controllers to sense chemical information and translate it into material activation and actuation. We conclude the dissertation with a discussion of possible future work and applications of the studies presented thereafter in *Chapter 6*.

1.5 Contributions

Chapter 2

A version of *Chapter 2* was published:

> Reprinted with permission from Joshua Fern, Jennifer Lu, and Rebecca Schulman. "The energy landscape for the self-assembly of a two-dimensional DNA origami complex." *ACS Nano*, 10(2), 1836-1844 (2016). Copyright 2016 American Chemical Society.

Joshua and Rebecca designed the research; Joshua and Jennifer performed the research; Joshua, Jennifer, and Rebecca analyzed data; Joshua and Rebecca wrote the paper.

Chapter 3

A version of *Chapter 3* was published:

> Reprinted with permission from Joshua Fern, Dominic Scalise, Angelo Cangialosi, Dylan Howie, Leo Potters, and Rebecca Schulman. "DNA strand-displacement timer circuits." *ACS Synthetic Biology*, 6(2), 190-193 (2017). Copyright 2017 American Chemical Society.

Joshua, Dominic, and Rebecca designed the research; Dominic conceptualized the circuit; Joshua and Dominic designed the circuit; Joshua, Angelo, Dylan, and Leo performed the research; Joshua performed the simulations and data analysis; Joshua, Dominic, and Rebecca wrote the paper.

Chapter 4

A version of *Chapter 4* was published:

> Reprinted with permission from Joshua Fern and Rebecca Schulman. "Design and characterization of DNA strand-displacement circuits in serum-supplemented cell medium." *ACS Synthetic Biology* (2017). Copyright 2017 American Chemical Society.

Joshua and Rebecca designed the research; Joshua performed the research; Joshua designed and performed the simulations; Joshua and Rebecca wrote the paper.

Chapter 5

A version of *Chapter 5* is in preparation for submission:

> Joshua Fern and Rebecca Schulman. "Modular DNA strand-displacement controllers that direct material expansion." *in preparation*.

Joshua and Rebecca designed the research; Joshua performed the research; Joshua and Rebecca wrote the paper.

CHAPTER 2: THE ENERGY LANDSCAPE FOR THE SELF-ASSEMBLY OF A TWO-DIMENSIONAL DNA ORIGAMI COMPLEX

Summary

While the self-assembly of different types of DNA origami into well-defined complexes could produce nanostructures on which thousands of locations can be independently functionalized with nanometer-scale precision, current assembly processes have low yields. Biomolecular complex formation requires relatively strong interactions and reversible assembly pathways that prevent kinetic trapping. To characterize how these issues control origami complex yields, the equilibrium constants for each possible reaction for the assembly of a heterotetrameric ring, the unit cell of a rectangular lattice, were measured using fluorescence colocalization microscopy. We found that origami interface structure controlled reaction free energies. Cooperativity, measured for the first time for a DNA nanostructure assembly reaction, was weak. Simulations of assembly kinetics suggest assembly occurs *via* parallel pathways with the primary mechanism of assembly being hierarchical: two dimers form that then bind to one another to complete the ring.

2.1 Introduction

Methods from DNA nanotechnology offer a programmable, inexpensive way to control the structure of matter at the length scale of a few to several hundred nanometers.[7] DNA nanostructures can be applied to control the spatial configuration of enzyme cascades,[24,25] metallic nanoparticles[36] and quantum dots.[157] Furthermore, these methods can be used to construct materials with tailored optical behaviors[33–35] or as tools for cell biology research[27,28] and super-resolution microscopy.[20,21]

In many cases, the application requires the assembled structure to have a particular size and geometry such that the molecules to be positioned at each of the many potential sites on an assembled structure can be independently controlled by the binding of functionalized DNA strands with different sequences. While DNA origami enable this control,[24,25] the length scale of assembled structures with independently addressable locations across the structure is currently limited to structures only slightly larger than individual DNA origami complexes.[77,78,158] While somewhat larger origami can be assembled, sequence costs increase linearly with structure area.[45] Techniques such as algorithmic self-assembly make it possible to build larger structures by reusing components, lowering these costs, but impart restrictions on the patterns that are practical to assemble and errors during assembly are frequent.[74,75]

An alternative to these methods is the hierarchical assembly of origami complexes from a heterogeneous mixture of origami monomers. In such a scheme, each origami component is assembled using a largely uniform set of staple sequences and each presents a distinct set of DNA sequence-dependent interfaces that control their interactions with other components. A different origami component assembles into each location of the final structure, so that each functionalizable site in the final larger structure is uniquely addressable.

The assembly of complexes from different types of components requires control over binding rates in order to achieve high yields.[159] In contrast to the assembly of regular lattices, the assembly of defined complexes requires not only reversible interactions between assembly components, but also a set of interfaces that minimize nonspecific binding between many pairs of components that should not bind.[79] Further, the reactions involved in forming the complex must both minimize kinetic traps and stabilize the completed structure.[65,67,69]

A quantitative understanding of the kinetics and energetics of component-component interactions involved in an origami self-assembly process would make it possible to systematically evaluate which factors limit yields and to then improve these yields through component and reaction process design. Previous studies have examined the kinetics and thermodynamics of the hierarchical assembly of DNA nanostructures using FRET or atomic force microscopy, but have been limited to origami 1D chain growth[78] or dimer assembly of significantly smaller nanostructures.[83,160,161]

Here we develop a novel assay to systematically measure the complete energy landscape of a heterogeneous self-assembly process for simple origami complexes and use it to understand Watson-Crick mediated interactions between origami components. This assay uses fluorescence microscopy and fluorophore colocalization, similar to methods previously used to detect the locations and contents of protein complexes within cells,[162] to quantify the concentrations of monomers, the many possible assembly intermediates and the target complex as an assembly process approaches equilibrium. We use this information to determine the equilibrium constants of assembly between different components and assembly intermediates, and build an energy landscape for the assembly process.

Energy landscapes are commonly used to understand how complex chemical reactions and assembly processes occur.[163] They can also be used to identify bottlenecks in supramolecular assembly[164–166] and explain how the assembly process varies under different physical conditions. Using information from these analyses, bottlenecks can be systematically overcome by changing the reaction conditions[167,168] or by designing new interactions between components or new components.[75,169,170] A better understanding of the microscopic rates of the assembly reactions could thus help optimize the assembly of origami lattice structures and, if applied to an assembly process involving different component types, could also be important for processes where rates are critical for controlling what structures are assembled.[65,68,69]

In the system we consider, the interfaces of four origami components present edges with pairs of complementary ssDNA overhangs. There are four binding interfaces that each have unique sets of sticky-end sequences; hybridization of all four interfaces produces an origami "ring" in which each origami structure is linked to two others along different interfaces. We used the assay we developed to measure the equilibrium constants for each possible combination of assembly reactants, including all combinations of origami monomer and assembly intermediates. These results suggest several important principles for reactions between origami components. In our system, origami preferentially bind along interfaces with similar structure, and the energy of interaction is not correlated with the hybridization energy of the single-stranded sticky-ends. We also find that the binding energy between two origami components does not appear to be affected by origami bound at other interfaces distant from the reaction site. Finally, we are able to directly measure the binding energy involved in ring closure, which involves reactants interacting simultaneously at multiple interface sites, and find that this set of reactions exhibit weak cooperativity. This cooperative interaction determines whether and how origami structures will demonstrate a preference for higher order structures in which most components are bound to multiple other components in the target structure.

For closed complex structures, the degree of cooperativity also determines whether complete complexes are energetically favored over assembly intermediates at equilibrium.

In addition, we use the information we collect about the different assembly reactions to produce a holistic view of the assembly process. By using simulations to estimate the total amount of material produced by each reaction over time during a typical assembly process, we found that assembly of tetramer complexes most likely occurred through pathways involving the initial formation of two dimers with the lowest $\Delta G°$, followed by either the assembly of these two dimer types or the sequential addition of monomers to one of these dimers. The method we lay out using the interfaces in this paper as an example structure therefore provides a foundation for analyzing the role of component design in heterogeneous assembly processes.

2.2 Materials and Methods

2.2.1 Origami design

The components of the self-assembly process were four two-dimensional rectangle origami, each composed of 191 staple strands with sequence complementarity to bacteria-phage M13mp18 DNA (Figure 2.1, Fig. S2.1–S2.3). Each origami was 32 helices wide and 219 bases long giving a size of about 101 x 75 nm as measured by atomic force microscopy. The helical twist of the origami structure was set at 10.43 bp/turn using selective deletion every 48 bases and by offsetting staple termination points along the length of the origami.[78] DNA hairpins inserted in the staple sequences in patterns of a "1", "2", "3," or "4" on the origami were used as distinguishing markers and visualized using atomic force microscopy. Origami used in fluorescence microscopy experiments also contained hairpins. Five staple strands, spaced across the origami, contained an additional docking region sequence on which a biotinylated DNA strand hybridized. These biotinylated strands enabled origami assemblies to attach to a Neutravidin coated glass surface. Eight other strands hybridized to the part of the M13 not complementary to staple strands and provided docking sites for fluorophore-labeled strands (Fig. S2.7). Staples along the edges of the origami were classified as either “blockers” or “linkers.” A blocker edge staple had a 4 nucleotide poly-T sequence extending from the origami edge designed to minimize nonspecific interactions between origami and inhibit excess linker staples in solution from binding to the origami and acting as additional sites for mediating origami–origami interaction (Figure 2.1B, Fig. S2.3). Linker staples produced an overhang on the origami components consisting of an eight-base region in which the two regions of the same strand hybridized with one another followed by five nucleotides complementary to the same region on one other linker. A specific interface for origami–origami binding was comprised of four linker and three blocker staples. On each origami, linkers extended from two of the four possible interfaces (Figure 2.1A–B) and were designated as interface “A” (top-left side), “B” (top-right side) or interface “D” (bottom-right side). Interface “C” was not involved in linking on any of the components and only contained blocker edge staples.

2.2.2 Origami preparation

DNA solutions for each origami were prepared as mixtures of scaffold strands (M13mp18, Bayou Biolabs), staple strands and labeling strands (Integrated DNA Technologies, Inc.). Staple strands consisted of structural, edge (linkers and blockers), fluorophore docking, and anchor docking strands. DNA strands containing either a fluorophore (ATTO647N, Cy3, or ATTO488) or a biotin molecule hybridized to docking sites on the origami. Origami that were not fluorescently labeled were prepared with a DNA strand that lacked a fluorophore but had the same sequence as the fluorophore labeling strands. These additional strands prevented excess, free fluorescently labeled strands from attaching to unlabeled origami. Hereafter, origami are noted by their number after a T (for origami tile) and the color of their fluorescent label, *e.g.*, T1-R for an ATTO647N labeled Tile-1 origami. Unlabeled origami have an "N" label, *e.g.*, T1-N. Origami stock solutions were prepared as 20 nM scaffold strand, 200 nM structural and anchor docking staples, 60 nM linkers, 200 nM blockers, 1200 nM biotinylated strands, 60 nM fluorophore docking strands, and 600 nM fluorophore-labeled strands. Before assembly, each monomer origami was prepared by annealing all above listed strands in TAE/Mg^{2+} buffer (40 mM Tris-acetate, 1 mM EDTA, supplemented with 12.5 mM magnesium acetate) from 90 to 20 °C at 1 °C/min using an Eppendorf Nexus thermocycler. Origami stock solutions were used without purification. AFM scans showed that <4% of all origami monomers were malformed (*e.g.*, broken or hairpins were indiscernible).

2.2.3 Assembly of origami into complexes and reaction equilibration

To assemble origami complexes, origami monomers were mixed in equal volumes from stock solutions to produce a mixture containing 5 nM of each origami type. This mixture was separated into aliquots, one per sampling temperature, and loaded into a thermocycler. During self-assembly, these mixtures were first heated to 55 °C for 20 min to melt any origami complexes (and not origami components themselves, see Section 2.5.1) that might have formed during the initial mixing of the origami monomers at room temperature. After 20 min, the mixtures were cooled from 55 to 25 °C at 15 min/°C, a rate at which origami–origami binding reactions were expected to equilibrate at each temperature step (Section 2.5.2).

2.2.4 Origami assembly reaction sampling

In increments of 5 °C between 50 and 20 °C, the composition of the reaction was measured using fluorescence microscopy. Fifteen minutes prior to sampling, tubes containing 200 μL of TAE/Mg^{2+} were added to a dry heater at the sampling temperature; the contents of these tubes were used to dilute the reaction samples. After quickly transferring individual reaction tubes to the dry heater, the solutions in the tubes were diluted by mixing 1 μL of the reaction solution with the dilution tube in the dry heater. For glass bottom dishes, 20 μL of diluted solution was added quickly after mixing to the Neutravidin treated surface containing 100 μL of TAE/Mg^{2+} and incubated for 3 seconds. Fifty microliters of diluted solution was incubated for 8 seconds on coverslips. After incubation, surfaces were quickly washed with TAE/Mg^{2+} to remove unbound origami and stored in TAE/Mg^{2+} until imaging. The attachment of origami using biotin-Neutravidin linkages ensured that origami could not travel on the surface or return to solution, so the structure of the monomers and complexes were expected to remain stable between sample preparation and imaging a few minutes to hours later. Fluorescent imaging was conducted on an Olympus IX71 inverted microscope with 1.6x magnification using a 60x/1.45 NA oil immersion objective (96x total magnification) and a cooled Andor iXon3 CCD camera using custom image-capture software. Three filter sets (Table S2.6) were used to separately image origami labeled with ATTO647N (red, R), ATTO488 (blue, B) or Cy3 (green, G).

2.2.5 Preparation of surfaces for fluorescence microscopy

Glass coverslips (VWR) or glass bottom dishes (*In Vitro* Scientific) were cleaned *via* sonication with 10 w/v% NaOH for 25 min. After washing with Milli-Q pure water, surfaces were treated with 50 μL of 0.5 mg/mL Biotin-BSA (Sigma Aldrich) in TNT buffer (10 mM Tris-HCL, 0.1 M NaCl, 0.05% Tween-20 at pH 7.5) for 30 min. After washing three times with TNT, glass surfaces were treated with 50 μL of 0.5 mg/mL Neutravidin (Thermo Scientific) in TNT buffer for 15 min. Excess Neutravidin was removed by washing the surfaces four times with TAE/Mg^{2+} buffer. All surfaces were prepared and used for microscopy measurements on the same day.

2.2.6 Atomic force microscopy

For AFM imaging, sample tubes were removed from the thermocycler and transferred to a heated glove box using a dry heater (all set to the reaction temperature being characterized). Inside the glovebox, 10 μL of TAE/Mg^{2+} was added to a small, freshly cleaved mica puck. Four microliters of reaction sample solution was added to the puck and incubated for one minute followed by washing with TAE/Mg^{2+} four times to remove excess DNA strands and origami species. Finally, 60 μL of TAE/Mg^{2+} was added and the puck was transferred to the AFM. All imaging was conducted using a Bruker Dimension Icon in ScanAsyst mode with a sharp nitride lever tip (SNL-10, tip C, Bruker) cantilever under fluid conditions.

2.3 Results and Discussion

2.3.1 Four origami components assemble into the target complex *via* Watson-Crick hybridization of sticky-end overhangs.

To determine whether the origami components with linking interfaces (Figure 2.1A–B, Figs. S2.1–S2.5) interacted as designed, we annealed a mixture of four origami tile components from a starting temperature of 55 °C down to 25 °C at 15 min/°C. Atomic force microscopy (AFM) analysis of the mixture at 25 °C showed the presence of tetramer structures (Figure 2.1C), but much of the material remained as either monomers or dimers. Few trimers were observed (Fig. S2.6).

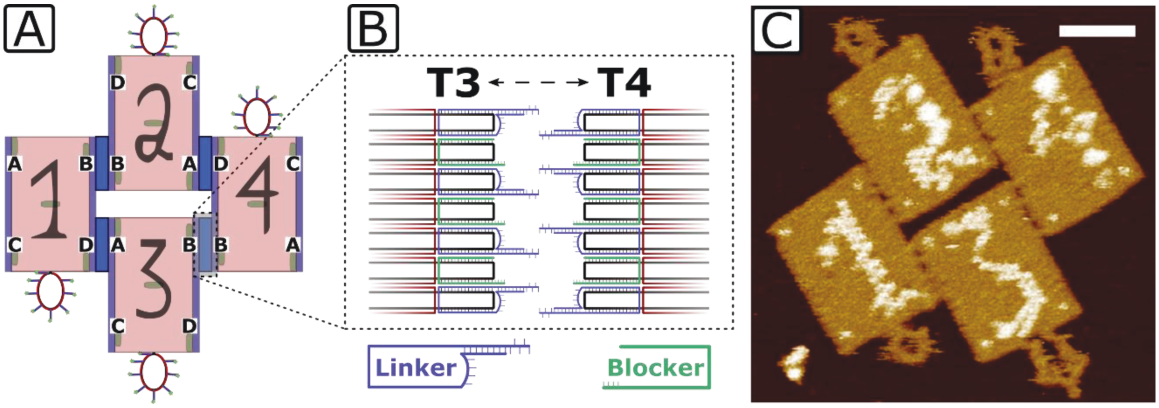


Figure 2.1: Multi-component origami assembly scheme. (A) Schematic of origami showing structural staples (red), hairpins on structural staples that produce distinguishing patterns on the AFM (black letters), edge staples that control interaction between origami components (blue regions), biotin-anchoring sites (green) and fluorophore-binding sites (loop-outs). Each origami binds another origami along one interface, coded by the letters A, B, C or D. (B) Edge staples act as either "blockers" (green, 4 thymine nucleotides) or "linkers" (blue, 8 bp dsDNA + 5 bp sticky-end) for origami-origami attachment (see Fig. S2.3). (C) AFM image of the target origami complex. In addition to the hairpins that produce visible numbers, the biotin-anchoring sites (corners and centers of each origami rectangle) and fluorophore docking sites (loops at top or bottom edge of origami) are visible. Scale bar 50 nm.

2.3.2 The concentrations of origami monomers and complexes after a reaction can be quantitatively measured using multicolor fluorescent labeling and fluorescence microscopy.

The first step in calculating the equilibrium energy landscape of multiple pathways in an assembly process is the determination of equilibrium constants for all possible reactions in the assembly process. These equilibrium constants can be calculated using the measured relative abundances of all possible species in the reaction. While AFM has been used to characterize the yields of origami self-assembly[60,72,78,171] because it enables clear visualization of the structure of assemblies, quantitatively measuring binding energies between species or the relative abundances of the different species is tedious because large surface areas must be scanned serially and the resulting images generally must be analyzed manually. Because the number of intermediate complexes scales exponentially with product complex size, the amount of AFM imaging required to accurately measure the abundance of all possible complexes quickly becomes prohibitive. To characterize the assembly landscape of our reaction without an imaging bottleneck, we therefore developed a method of measuring the relative abundances of DNA origami monomers, intermediates and target complexes using fluorescence microscopy and fluorescently tagged DNA origami (Fig. S2.7–S2.8). Because this method enables the identification of hundreds of origami monomers and complexes using a single epi-fluorescent image capture, it makes it possible to rapidly and accurately measure the fraction of material of many different species whose abundance may vary. To distinguish different origami species, each of the four origami were labeled with either ATTO647N (R), Cy3 (G) or ATTO488 (B) fluorophores or kept unlabeled in a given reaction. The composition of the complex could then be determined by the combination of fluorophores it possesses.

One potential source of error in our fluorescence measurements is that when the density of origami on the glass surface is high, components can land near one another by chance so that they appear bound to one another. We used simulations to determine how often these events would occur and the amount it would skew our measurements of complex abundances using fluorescence colocalization (Figs. S2.9–S2.10). We found that for typical densities of origami components on slides, about 3% of origami components that were unbound appeared to be in assemblies.

To verify that our fluorescence co-localization microscopy technique could be used to characterize the relative abundances of origami and origami complexes, we compared the results of the technique with measurements of the abundances of complexes taken using atomic force microscopy. We mixed sets of two origami that were designed to bind to one another, heated them to 55 °C and then slowly cooled them using an annealing schedule that ensured the reaction had approached equilibrium between 50 and 25 ºC (Section 2.5.1–2.5.2). Each of the two origami monomers were labeled with a different fluorophore so that static fluorescence micrographs of surface-bound origami provided a 2D representation of reaction composition (Figure 2.2A). The relative fraction of species in each possible configuration (the two monomers or the dimer) was determined by counting the number of objects that appeared in one or two fluorophore channels (Figure 2.2B). We compared the reaction compositions of a two-component reaction mixture measured using our microscopy methods with those from the same reaction characterized using AFM scans and found that the measured fraction of material in a dimer complex measured using our fluorescence assay and AFM micrographs were similar to one another, suggesting that the fluorescence microscopy assay we developed can be used to quantitatively measure the fraction of material in different reaction states (Figs. S2.11–S2.13, Table S2.1).

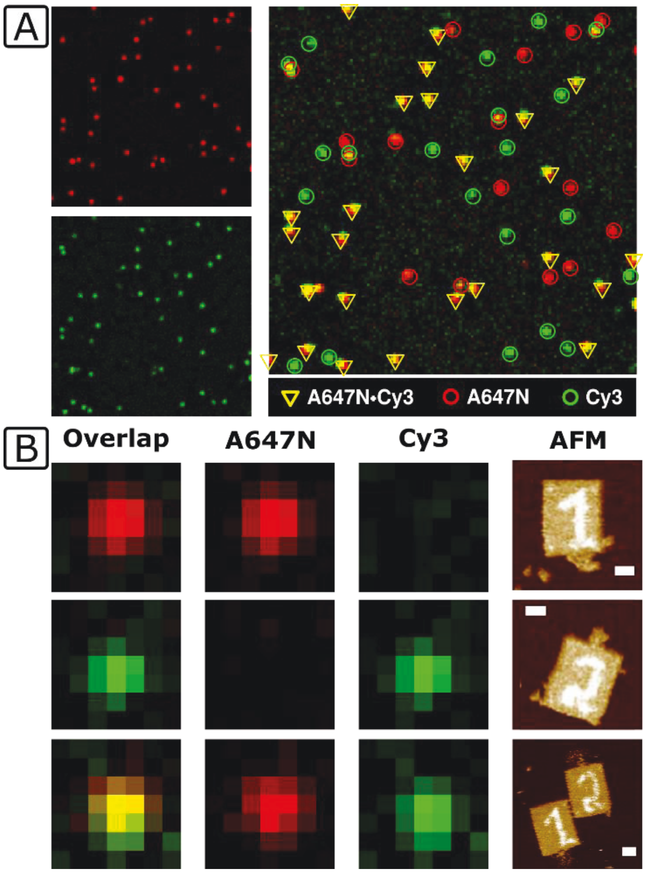


Figure 2.2: Multi-color fluorophore labeling of different origami components allows the assembly state of origami complexes to be determined from multi-color fluorescence micrographs. (A) The two origami tile types are labeled with two different fluorophore types for the assembly of T1-ATTO647N and T2-Cy3 into a T12 dimer. Pixel locations of apparent origami structures are visualized in individual fluorescence channels (red and green). Fluorescent objects that appear in multiple channels are origami assemblies (right image, labeled by colored shapes designating which channels). Section 2.5.3 details the algorithm used to detect and count fluorescent objects. Image sizes: 21.4x21.4 microns. (B) Example fluorescent origami assembly species, both in individual channels and overlays, for objects in one or two channels and an AFM representation of the specified fluorescent object. Size of each fluorescent image: 1.2x1.2 microns; AFM scale bars: 25 nm.

After verifying that our assay could be used to measure the fraction of material in different configurations close to equilibrium, we used it to find equilibrium yields in all possible reactions involving two or three binding species. Since there was a 1:1 ratio between the possible number of fluorophore types (3) and the number of origami components (up to 3 in dimerization and trimerization reaction mixtures), the fraction of material values for all species in a reaction were directly calculated from counts of objects in the fluorescent images (Fig. S2.14, Figure 2.3A).

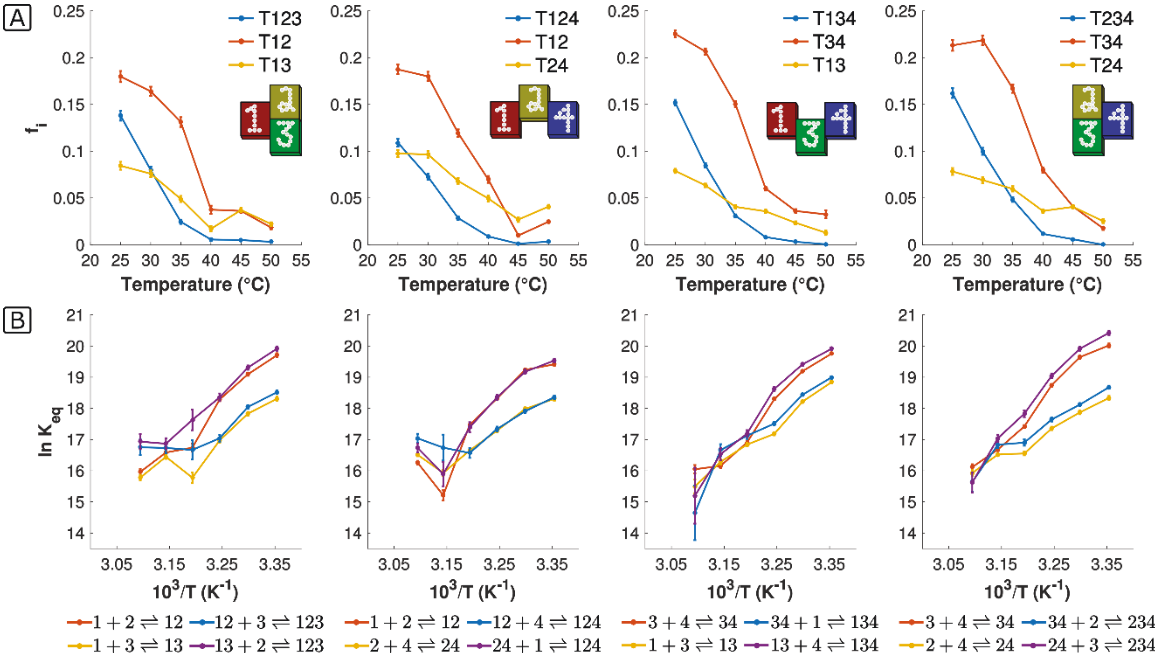


Figure 2.3: The distribution of assembly material among potential species in reactions involving 3 origami components close to thermodynamic equilibrium at a range of temperatures and the corresponding equilibrium constants for the assembly reactions. (A) The fraction of assembly material (f_i) in four reaction mixtures involving three of the four origami components (left to right): 1-2-3, 1-2-4, 1-3-4 and 2-3-4. (B) Van't Hoff plots (log of the equilibrium constant *vs.* inverse temperature) for the possible reactions between the components for the reactions in (A). Equilibrium constants for reactions where binding occurs between the same types of interfaces (1-2 and 3-4) are larger than those for reactions where binding occurs between different interface types (1-3 and 2-4). Non-linear behavior in the Van't Hoff plots precluded the calculation of standard enthalpy and entropy changes for each assembly reaction. Here and elsewhere, error bars were calculated as described in Section 2.5.5.

2.3.3 The binding interfaces of origami play a role in the equilibrium constants of origami-origami interactions.

Previous experiments have shown qualitatively that the amount of binding between origami can be dependent on the structure of the origami components,[60] which can vary across a rectangular origami structure because of asymmetries in either crossover patterns or sequence.[172] The four interfaces of the origami components we designed, therefore, might be expected to bind to one another in reactions with different equilibrium constants, even if the sticky-end sequences that hybridized during binding were the same. In support of these ideas, we found that the K_{eq} of a reaction was dependent upon which interfaces of the origami were interacting (Figure 2.3B). Relatively strong binding was observed between origami monomer pairs T1-T2 and T3-T4, where binding occurs between two origami interfaces with identical structure - interface "B" (Figure 2.1A). Weaker interactions were observed between T1-T3 and T2-T4, where binding occurs between the dissimilar origami interfaces "D" and "A". In contrast, the hybridization energy for each interface's binding predicted by considering the hybridization of the sticky-end sequences using the nearest-neighbor model[173] did not correlate with the measured equilibrium binding constants for origami binding, as free energies of binding predicted using this method were very close to equal (Table S2.2).

In protein assembly, allosteric interactions between components are common. Allostery implies that when two components bind, the conformations of the components change, altering the binding energy between these components and future binding partners. To determine whether allosteric interactions were present in the assembly of origami complexes, presumably because of conformational changes incurred by the origami as they hybridize, we compared equilibrium constants of a dimerization reaction in the presence or absence of a third origami component capable of binding a secondary interface. For example, in the

T1+T2, T2+T4 and T1+T2+T4 reaction systems, we found that the addition of a third component to a reaction mixture did not have a significant effect on the K_{eq} of a reaction, implying that allosteric interactions do not have a major role in origami–origami binding here (Figs. S2.16–S2.17).

2.3.4 Equilibrium analysis of four-component systems.

The number of unique fluorophore types available to a conventional microscopy system is often limited to three or four due to the spectral overlaps between fluorophores and limitations in available light filter sets. This limitation is an obstacle to the use of the described assay for the analysis of reaction mixtures with more components than fluorophore types. To scale our technique for measuring the equilibrium constants of individual reactions to systems with more than three components, we developed a method to measure the contents of reaction mixtures by running the same reaction several times using different labeling schemes for the components in each reaction, and combining the data from these reactions to find the fraction of material in each possible monomer or complex. Because the fluorophore labels on the components did not affect their binding behavior (Fig. S2.11), we assumed that the equilibrium constants in each of the reaction mixtures were the same. To study how the fraction of assembly material was distributed among the possible species in a reaction to form a heterotetrameric ring (Figure 2.1), we performed three assembly reactions and measured the distribution of colocalized fluorophores in each reaction mixture. In two out of the three reaction mixtures, three of the origami were labeled with three different types of fluorophores and a fourth origami was not labeled (unlabeled origami were denoted with an "N"). In the third reaction mixture, the fourth origami was labeled with the same type of fluorophore as one of the other origami (Figure 2.4A).

We counted the abundance of each of the colocalized fluorescent objects in each experiment in the same manner as in the two- and three-component mixtures (*e.g.*, the number of objects in each fluorescent channel at the same location in each of the labeling schemes). The abundances of each possible combination of colocalized fluorescent labels in each of the three experiments were then related by mass balance equations describing the possible fluorescent objects representing a given assembly species (Table S2.3). For example, in the T1234-RNGB scheme, an object that fluoresces in the red and green channels represents either a T13 or a T123 species due to the "2" origami being unlabeled, so the fraction of material that fluoresces in red and green is the sum of the fractions of the material that are either T13 or T123. Assuming that the fraction of material of each species does not change between reaction mixtures, the mixtures with different labels will each constrain our knowledge about the values of these abundances, making it possible to calculate them given all of the information from the three reactions with different fluorophore labeling schemes. We used a least-squares solver using the mass balance equations to find the best fit to the fraction of material of the possible species, with the bounding constraints that the fraction of material of each species must be nonnegative (Section 2.5.7). However, we found that this method for determining the fraction of material and equilibrium constants was difficult numerically: the fits to the fraction of material for each species were highly dependent on the assumption that the concentration of each origami component was exactly the same in each experiment because, in practice, the concentration of the origami components varied slightly between the reaction mixtures because of effects such as pipetting error (Fig. S2.18).

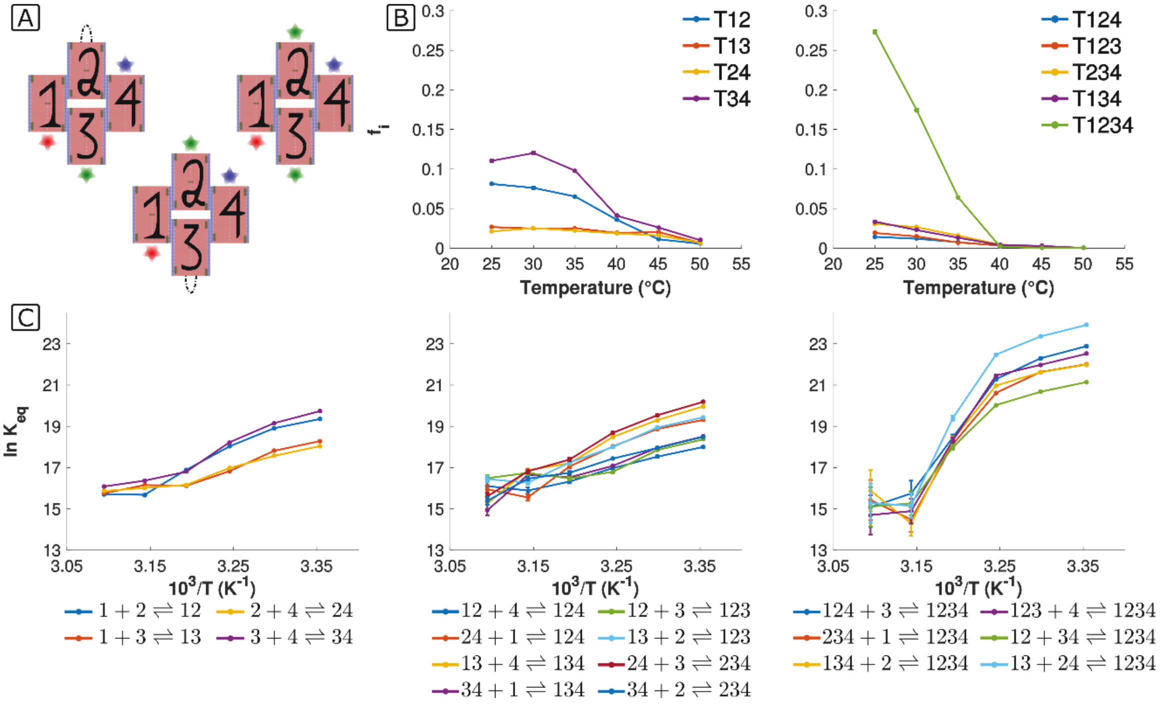


Figure 2.4: Determination of the equilibrium constants for a four-component origami assembly reaction. (A) Example set of origami component labeling schemes used when measuring the distribution of complexes in four-component assembly: either T2, T3 or both T2 and T3 were labeled with a Cy3 fluorophore in three separate experiments (these labeling schemes were termed T1234-RNGB/RGNB/RGGB respectively). We also characterized the four component assembly process using the labeling schemes T1234-NRBG/GRBN/GRBG. Comparison of the results predicted using these two labeled schemes were also used identify and correct numerical instabilities in our analysis process (Sections 2.5.7–2.5.8). (B) Combining the fluorescent object abundances from each reaction mixture enabled the calculation of the fraction of reaction material (f_i) of all fifteen possible species (See text). At 25 °C, the majority of components existed as monomers (~33%) or tetramers (~30%). (C) Equilibrium constants for the eighteen assembly reactions. Tetramerization reactions showed higher equilibrium constants than dimerization and trimerization reactions at every measured temperature, suggesting the presence of cooperative interactions. Data in (B) and (C) were determined using both the RNGB/RGNB/RGGB and NRBG/GRBN/GRBG labeling schemes (see Section 2.5.5).

To estimate the concentrations of the different species in a manner that was less sensitive to slight measurement variations between experiments, we modified the system of equations used in our analysis. On the basis of our observations that reactions between origami components do not exhibit allostery, we assumed that the equilibrium constants for the dimerization and trimerization reactions in the four-component mixture were close to those for the reactions measured in the two- and three-component mixtures. We used this assumption to constrain the equilibrium constants for the dimerization and trimerization reactions to be close to those we had previously measured. The resulting fits produced the same results for different subsets of experiments, indicating that they were not as sensitive to small variations between different assembly reactions (Section 2.5.8, Fig. S2.19). The resulting measured fractions of material and equilibrium constants are shown in Figure 2.4B, C.

2.3.5 Ring closure reactions exhibit cooperative interactions with equilibrium constants controlled by interface pairings.

Measuring the equilibrium constants for the reactions that form four-component origami complexes allowed us to determine the binding energy for reactions that formed the target complex (Figure 2.4C). These reactions, which included those between a monomer and trimer and those between two origami dimers, allowed two origami interfaces to interact simultaneously. At every temperature, the K_{eq} for each of these reactions was larger than the K_{eq} measured for any dimerization and trimerization reaction, suggesting the presence of cooperative interactions. That is, the reaction of the components by multiple origami binding interfaces increased the energy of interaction (lower $\Delta G°$). Because of the symmetry in the design of the origami complex assembly, all of the trimer-monomer reactions involved both B-B and A-D interface binding (*e.g.*, T123+T4 reaction has T3-T4 and T2-T4 interface binding reactions) and resulted in equally balanced K_{eq} values. However, the reactions between two dimers consisted of either two B–B interface reactions (T13+T24) or two A–D interface reactions (T12+T34). With these reactions, there was a marked difference in equilibrium constants; the B–B interface reaction had about a 20-fold higher K_{eq} than the A–D interface reactions. This followed the pattern observed earlier in dimerization and trimerization reactions where B–B interface reactions had a greater equilibrium constant in the studied system as compared with A–D interface binding reactions.

2.3.6 Reaction equilibrium constants provide insight into potential asymmetries in assembly pathways.

We used the equilibrium constants for all possible reactions in the four-origami component system and kinetic assembly simulations to produce an energy landscape for assembly. This landscape shows which relative pathways from origami monomers to the target complex were the most energetically favorable (Figure 2.5). The energy landscape is important because its shape can determine the yield of assembly;[163,164] in particular, one potential principle for design is the idea that a biased assembly pathway, in which a single assembly process tends to be followed, is an efficient method for self-assembling complexes.[65,66]

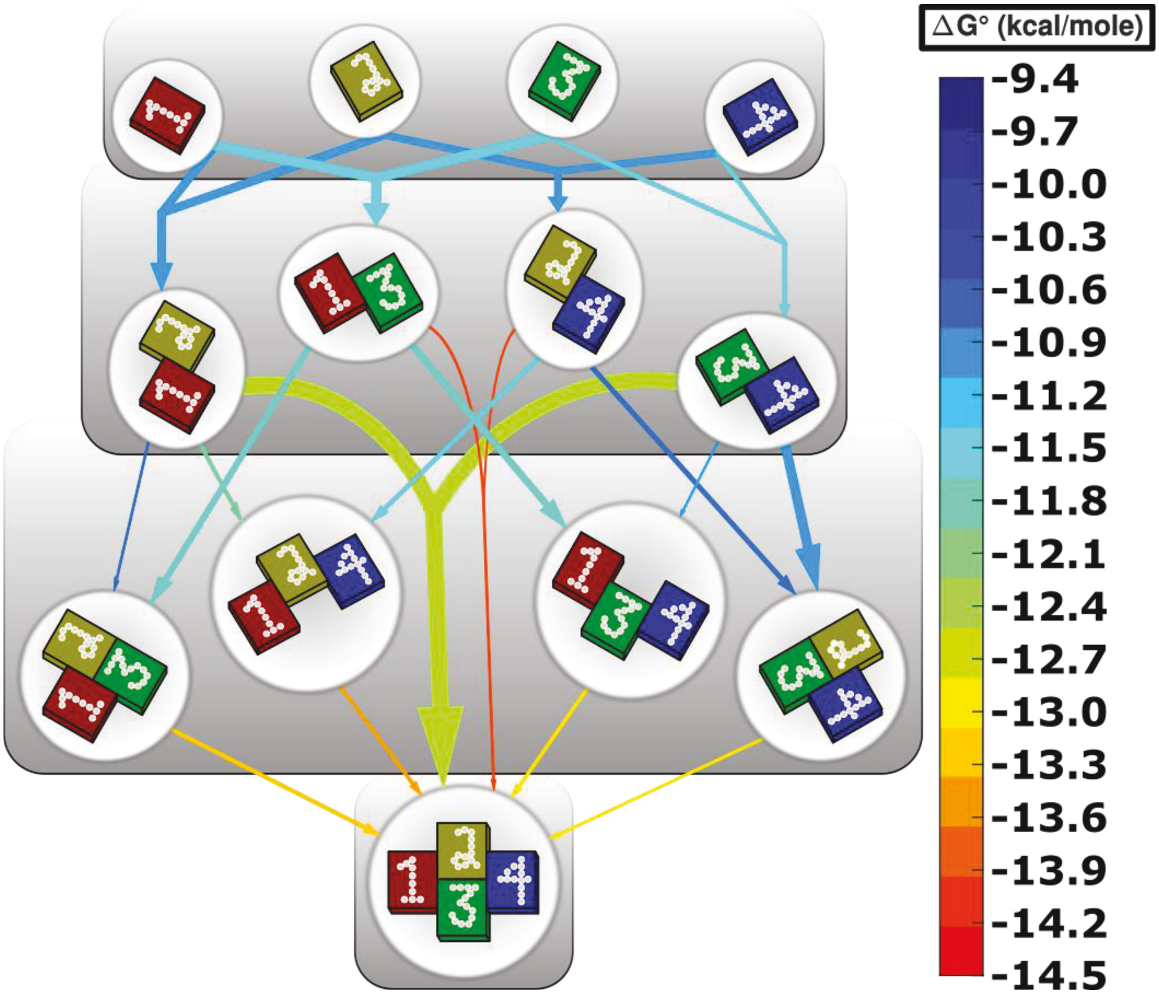


Figure 2.5: The assembly landscape for a four-component origami assembly reaction. Pathways in the self-assembly of a four-component nanostructure are shown for assembly at 25 °C. Each arrow represents a possible reaction between components with the product complex at the arrow's end. The line color is the calculated free energy change from experimental measurements using $\Delta G°=-RT\ln K_{eq}$ (from blue to red is more negative) and the width of the line indicates the relative amount of material generated by each reaction in kinetic simulations. For clarity, arrows for monomers are omitted for reactions involving a monomer and a larger assembly.

We created a simple kinetic simulation of the assembly process to understand how assembly of tetramer complexes might occur. Following models and experiments analyzing DNA tile crystallization[74] and DNA origami oligomerization,[171] we considered how assembly would occur if the system had a single on-rate for all reactions. Using this on-rate, we then calculated the off-rate for each reaction using the equilibrium constants that we experimentally measured, and used these rates in a mass-action kinetic simulation of the assembly process of the four origami components (Section 2.5.9, Figs. S2.20–S2.22).

We calculated the relative amount of material produced by each reaction in our simulations. For dimerization reactions, the major products generated were T12 and T34, and because the concentrations of these two dimer species were much higher than the other two dimers, the reaction rates for subsequent reactions from T12 or T34, either when adding a third monomer or the T12+T34 dimer–dimer reaction, were also large, despite those interfaces having lower binding energies. The most prevalent tetramerization reaction in our simulations was T12+T34, forming about 2x more tetramers at 25 °C than the next most favored reaction, T234+T1 (Fig. S2.22). While T12+T34 was the dominant tetramerization reaction at temperatures less than 40 °C, the dominant assembly pathway varied with temperature (Fig. S2.22). An energy diagram, depicting the free energy and the relative amount of produced complex for each reaction at 25 °C is shown in Figure 2.5.

2.4 Conclusion

In this paper, we develop a method using standard fluorescence microscopy techniques to determine the equilibrium constants for the assembly of heterogeneous DNA origami complexes. The techniques outlined are straightforward for the complete analysis of small assembly reactions in which each component can be labeled with a different fluorophore type, and can also be extended to include more components by repeating reactions with different labeling combinations, and by using mass balance equations and the equilibrium constants measured in simpler reactions. These techniques could also be used to characterize the thermodynamics of 3-dimensional origami assembly processes.

The energy landscape we determined demonstrates important principles that govern the energetics of origami complex assembly, and can be used to increase the yield and efficiency of self-assembly. We found that the energy of interaction between different components appeared to be controlled not by the specific sequences in the Watson-Crick binding pairs, but by the structure of the origami interface. This result suggests that, despite the importance of sequence effects in DNA hybridization, in the assembly of larger origami structures the structure of the interfaces largely determines binding energies, even when the binding process is actually mediated by DNA hybridization. These results expand upon studies of smaller DNA nanostructures where the architecture of small DNA tiles impacts the thermodynamics and kinetics of dimerization.[161,174] It would therefore be interesting to characterize how the structure (*e.g.*, rigidity) of origami and their interfaces affect the on-rate of origami–origami binding.

Understanding how to assemble DNA origami structures using Watson-Crick base-pairing is important because, while blunt-end base stacking is an alternative method to hierarchically self-assemble origami into larger structures, there are a relatively small number of distinct blunt-end stacking interfaces[78] and the resulting structures would be difficult to stabilize *via* processes such as enzymatic ligation.

Additionally, explicit consideration of the structure of origami components at the nanometer scale as predicted in simulation[82,175] or experiment[176] could be used to build higher-yield tunable interfaces between origami components.[80] The fact that binding energies are so dependent on structure may also make it possible to use these sorts of measurements to quickly and inexpensively test hypotheses about the average structure and the ensemble of structures of particular origami assemblies.[82]

Our techniques further suggest a way to explicitly measure cooperative interactions between multiple component types. In the case of homogeneous assembly, these methods would make it possible to measure cooperativity easily to aid in the construction of large lattices. Some degree of cooperativity is required for the assembly of regular crystals with few defects[74,76] and also, in the case of closed complex structures like those studied here, controls the stability of the final complex. Cooperativity is also important for the control of the self-assembled product in processes such as algorithmic self-assembly.[74]

Finally, the technique developed here allows the elucidation of the complete energy landscape for assembly and therefore can be important for a systems-level analysis of origami component assembly processes. While the development of principles for the self-assembly of complexes is still being elucidated,[65,67–69] it is clear that the ensemble of binding energies and how they are distributed between reactions is a major determinant of assembly speed and yield. The ability to measure all of the binding energies for a single assembly process, and perhaps to also tune these energies by designing alternative pairs of interfaces, will be important for a rational, reliable route to the self-assembly of origami complexes and more generally, to the hierarchical self-assembly of DNA structures that bridge different length scales.

2.5 Supplemental Information

2.5.1. Initial melting of origami complexes

After mixing the origami components, the reaction solutions were heated to 55 °C for 20 minutes to melt any origami complexes that might have formed during the initial monomer mixing. Atomic force microscopy verified that this heating process did not melt the monomer origami: specifically, we observed no incidences of hairpin patterns interchanging between origami tiles after the heating and cooling process (Fig. 2.1C, Fig. S2.6). Other experiments have shown that the probability of a linker edge staple displacing a blocker when heating to 55 °C from 20 °C is less than 5% and a blocker edge staple displacing a linker staple is less than 1%, indicating that the displacement or melting of edge staples had a minimal impact on experiments.[44]

2.5.2 Equilibrium assembly of origami complexes

Origami complexes were assembled by cooling a mixture of multiple origami components from 55 to 25 °C at a rate of 15 minutes/°C. For dimerization reactions, this annealing schedule is compatible with recently measured rates of reaction and rapid equilibration for the hybridization of two origami components.[44] To check that the same annealing schedule permitted reactions in four-component reaction mixtures to reach equilibrium, we measured the fraction of reaction material in each assembly state for both a cooling and a re-heating cycle using the T1234-RNGB, RGNB and RGGB labeling schemes (results shown in Fig. S2.23, analysis as described in Section 2.5.8). The reaction solutions were held at 25 °C overnight in between annealing and re-heating. The match between cooling and re-heating measurements indicate that the cooling rate of 15 minutes/°C is also slow enough for reactions in the assembly of the target tetramer complex to equilibrate.

2.5.3 Image analysis and derivation of assembly compositions

After image acquisition, fluorescent object locations were determined using standard object detection algorithms written in MATLAB (source code available upon request). Specifically, images were first subjected to a Gaussian filter of size 100 and a sigma value of 5 for background subtraction. A threshold to generate a binary mask of the image then required potential objects to have an intensity 5% greater than the local background intensity and less than 50x the local background (to remove large fluorescent objects). Finally, because the extent of the detected features varied in practice, if two objects in the same channel were found to be in neighboring pixels, the object with the highest intensity was chosen as the true object. Each object's intensity was approximately 1.2-3x above the object's local background intensity, where the local area is defined as the 3x3 pixel area around a fluorescent object (Fig. S2.8). Object locations were compared across fluorophore channels to determine if an object consisted of single or multiple fluorophore types. Objects were considered to be present in more than one color channel if their pixel locations resulted in a Euclidean distance less than √2 pixels, approximately 236 nm. The maximum theoretical distance between fluorophores on different origami was roughly 200 nm. The total number of overlapping fluorescent objects was determined for each subset of fluorescent channels (red alone, red+blue, red+green, red+green+blue, etc.). For each single origami, assembly intermediate and the target complex, the fraction of reaction material in configuration *i* (f_i) was calculated using the formula:

$$f_i = \frac{\text{Total number of origami in assembly species } i}{\text{Total number of origami in all assembly species}} = \frac{\upsilon_i N_i}{\sum \upsilon_j N_j} \qquad (S2.1)$$

where N_i is the number of the detected fluorescent objects of assembly species *i,* and υ_i is the number of origami tiles in assembly species *i*. The fraction of reaction material is a measure of how much material is present in a specific assembly species compared to the total material of the system.

2.5.4 Effect of fluorescent object surface density on apparent assembly abundances

Since the fraction of material in a given assembly state was determined using fluorophore colocalization, the density of fluorescent objects could affect the calculated number of colocalized objects. Specifically, because the origami bound to the surface in a random fashion (we did not direct where they bound), there was a non-zero probability of two (or more) fluorescent origami binding the surface such that the distance between them did not allow them to be resolved as separate species rather than an assembly complex. To estimate the effect of plating density on our measurement error due to this effect, we simulated a series of experiments involving different densities of objects that attached to the surface in random locations (Figures S2.5 and S2.6). Two cases were simulated: (1) all monomers were unbound, so the expected fraction of material in complex form was 0 (2) different fractions of the origami were in complexes. We expect a situation like case 1 above the melting temperature of the complexes (*e.g.*, 50 °C) and a situation like case 2 at intermediate to low temperatures where binding between origami can occur. In each simulation, we set a number of each monomer type (called the surface density) to be deposited on the surface (all monomer types were assumed to be present in equal number). The probability of binding a surface site was equal across the virtual surface. The simulated surface densities ranged from 50 to 2500 components per field of view, which is representative of the densities of origami we observed in our experiments. In the majority of experiments, prepared surfaces contained 800 – 1000 total components per field of view (512 x 512 pixels). Ten surfaces were seeded for each simulated surface density.

In the simulations for case 1, where no origami were bound to one another, all components were placed on a 512 x 512 grid by randomly selecting their X and Y coordinates. The fraction of material in each assembly state was then calculated from these X/Y coordinates using the same methods we used for analyzing experimental data. A small fraction of origami components (with the fraction being dependent on density) appeared bound to one another in two component simulations (Fig. S2.9). In simulations involving three component types that were not expected to bind to one another (Fig. S2.10A), our analysis also showed that a small fraction of components appeared to be part of complexes, and that the number of spurious detected complexes decreased as the size of the complex increased. This is logical as the probability of three objects (or four) landing on the same or neighboring pixels on a 512 x 512 grid is much lower than two objects.

For case 2 simulations, where some origami were actually bound to one another, we set the fraction of origami in different complexes to the fractions that we observed in a typical experiment (Figs. S2.9 and S2.10B). Complexes and monomers were randomly placed on a 512 x 512 grid as above. For complexes, two (or more) monomers were placed at the same X/Y coordinates as their binding partner(s). In this case, the plating density of a component was the total number of objects in complexes plus the number of single monomers. For example, a surface with 2 component types, 2000 components of each type and a dimer assembly abundance (f_i) of 20% contains 1600 of component 1, 1600 of component 2, and 400 dimers. The results of these simulations allowed us to estimate the amount of error in our measurements as a function of object density on the surface.

2.5.5 Estimation of uncertainty in experimental measurements and calculations

To estimate the uncertainty in our measured fractions of material in a given assembly species for an individual experiment, we employed standard methods for inferring an estimated proportion (here the fraction of material) *via* bootstrapping and the associated variance in each proportion.[177] For each assembly mixture, 5–10 locations on the dish or coverslip were imaged for each measured temperature point. The total number of detected colocalized objects for each reaction species was summed across the locations to produce a total count for each reaction species. For each reaction species (*e.g.*, T124-RGB), a random sample (of size *n*) equal to the total number of detected objects of that species type was drawn with replacement from the population of colocalized objects. The binomial probability (*p*) of an object being the specified species was calculated from this sample population. This procedure was then repeated for each reaction species. The probabilities (p_i) are related to the fraction of reaction material (f_i) in a given reaction state through the stoichiometry (v_i) on the number of origami tiles in said state (see Section 2.5.3):

$$f_i = \frac{v_i p_i}{\sum v_j p_j} \qquad (S2.2)$$

This sampling was repeated 1000x and the average fraction of material was calculated for each reaction species. The fraction of material for a reaction species is also the probability a fluorescently detected object on the surface is present in an individual fluorophore channel (Red, Green, or Blue object). For each f_i, the probability $q=1-f_i$ is the probability of a fluorescent object in the population of objects _**not**_ being a given species. The confidence intervals for each probability was then calculated from the following equation:[177]

$$\frac{f + \frac{z^2}{2n} - \sqrt{\frac{fqz^2}{n} + \frac{z^4}{4n^2}}}{1 + \frac{z^2}{n}} \leq f \leq \frac{f + \frac{z^2}{2n} + \sqrt{\frac{fqz^2}{n} + \frac{z^4}{4n^2}}}{1 + \frac{z^2}{n}} \qquad (S2.3)$$

where *z* is equal to 1.96 for 95% confidence intervals and *n* is the sample size of detected species. Figure S2.15 shows a sample workflow for this analysis. The average fractions of material and 95% confidence intervals reported in Figures 2.3–2.4 and in the Supplemental Figures below were calculated using the above method. Figures in the main section and in the supplemental info containing replicate experiments are represented by a weighted average of the replicate experiments with error bars calculated from a weighted variance.

The amount of error calculated using the above method represents the error associated with the amount of data collected about each species in a single experiment. These errors do not include errors due to experimental variation in the concentrations of the origami components that would be expected from reaction mixture to reaction mixture or other systematic error due to measurement technique/protocol limitations. To attempt to measure the error due to these various sources, we repeated a single experiment in triplicate and other experiments in duplicate (list in Table S2.5, comparison of experiments in Figure S2.25). For a mixture of 2 components, we found that the variance between experiments was within 0.2 % of the variance within a given experiment. For mixtures of 3 components, the variance between experiments was larger, likely due to an increased variance in a tiles' concentration when increasing the number of components (*i.e.*, the ratio of Red:Green:Blue detected objects in a population of objects). The effect of this source of error might be expected to increase as the number of monomer components in the target structure increases.

The mean and 95% confidence intervals of the calculated fractions of material for each species were used to calculate the mean and variance of the equilibrium constants for each reaction. Specifically, we generated a sample population for each fraction of material using Gaussian distributions with the means and 95% confidence intervals calculated above. For 1000 iterations, we sampled these fractions of material to generate a pool of equilibrium constants for each reaction. This pool of equilibrium constants was then used to calculate the mean and standard deviation for each equilibrium constant of each reaction.

2.5.6 Calculation of equilibrium constants from the fractions of reaction material

To calculate the equilibrium constants for each assembly reaction, we first derived an equation relating the fraction of reaction material in a given assembly species with its concentration. The concentration of a species can be written in terms of its fraction of reaction material using the equation:

$$C_i = \frac{f_i}{v_i}\sum v_j C_{0j} \qquad (S2.4)$$

where C_i is the concentration of assembly species *i* and C_o is the initial concentration of monomer species *j*. In the experiments presented here, all origami monomers had an initial concentration of 5 nM; the initial concentrations of other species were 0. The following equation was used to calculate the equilibrium constants for each reaction:

$$K_{eq} = \frac{C_P}{C_{R1}C_{R2}} = \left(\frac{f_P}{f_{R1}f_{R2}}\right)\left(\frac{v_{R1}v_{R2}}{v_P}\right)\left(\frac{1}{\sum v_j C_{0j}}\right) \qquad (S2.5)$$

where *P* indicates the product of a reaction where *R1* and *R2* are the two reactants.

2.5.7 Calculation of the fractions of reaction material during target complex assembly using mass balance equations alone

In our experiments determining the fraction of material assembled using all four components, only three types of fluorophores were used to label the origami. To calculate the fraction of material in each assembly state using fluorescent colocalization, fluorescent object abundances were measured for the same reaction using three different labeling schemes (*e.g.*, T1234-RGNB, RNGB and RGGB). The measurements from each of these reactions were combined to determine the fraction of material of each possible monomer type or complex after the reaction.

In two of the labeling schemes where one of the origami was unlabeled (indicated by an "N"), any detected fluorescent object could represent more than one assembly species (Table

S2.3). For example, the number of "RG" fluorescent species in the T1234-RGNB reaction mixture is the sum of the number of T12 and T123 in the mixture and is represented by the following equation

$$N_{RG} = n_{T12} + n_{T123} \tag{S2.6}$$

By switching which origami was unlabeled, three systems of mass balance equations relating fluorescent objects with which assembly specie(s) it could represent were generated. In each reaction mixture, there were 7 possible combinations of fluorescent objects (*e.g.*, red alone, red+blue, red+green+blue) giving a total of 21 mass balance equations. Thus, by measuring the relative abundance of each fluorescent object in each of the three schemes and using these mass balance equations, we were able to calculate the relative abundance of each assembly species in an assembly population based upon the three measured reaction mixtures.

We used a standard least-squares algorithm written in MATLAB to calculate the fraction of material in each assembly state in the assembly process. To facilitate logical solutions from the algorithm, we applied a lower bound of zero on allowed concentration solutions to ensure that only positive fractions of assembly species were calculated. The residuals to be minimized by the solver were calculated using the following equation:

$$\bar{r} = \bar{F} * \bar{a} - \bar{N} \tag{S2.7}$$

where $\bar{r}$ is a vector of residuals, $\bar{F}$ is the matrix pairing the fluorescent objects with which assembly species they could represent, $\bar{a}$ is a vector of the relative abundances of each assembly species (15 species total) and $\bar{N}$ is a vector of the relative abundance of each fluorescent object in each of the 3 labeling scheme populations. The solver finds the vector $\bar{a}$ that minimizes the sum of the squares of the residual vector $\bar{r}$.

To test whether this analysis method produced reliable results, we used it to find the fractions of material for each of the species predicted by two assembly experiments, each involving three labeling schemes. In the first set of experiments we labeled the 4 origami using the RNGB/RGNB/RGGB schemes. In the second set of experiments we labeled the 4 origami components using the NRBG/GRBN/GRBG schemes. Since the presence of fluorophore labels did not affect the fractions of material in dimer form (Fig. S2.11, Table S2.1), we would not expect the specific set of labels to affect the fractions of materials measured in other reactions. However, when we compared the predicted fractions from the two experiments we found significant differences in the calculated abundance of two trimer species (Fig. S2.18). While the results from the RNGB/RGNB/RGGB schemes showed a relatively high abundance of T123 and low abundance of T124, the NRBG/GRBN/GRBG schemes showed a relatively high abundance of T124 and low abundance of T123. In both of these sets of schemes, the trimer with high abundance was the complex containing one origami with ATTO647N labeling, and two origami that were either Cy3 labeled or unlabeled. However, the fraction of reaction material in each assembly state measured using AFM showed the abundance of each trimer species to be relatively equal and low in value (Fig. S2.24, Table S2.4). We hypothesized that the differences in the results were due to numerical instabilities in this analysis method, such that small amounts of noise in the measured numbers of different types of fluorescent objects could produce large changes in the predicted fractions of the material of the different species.

2.5.8 Calculation of the fractions of reaction material during target complex assembly using mass balance equations and the equilibrium constants for reactions producing assembly intermediates

To prevent the numerical instabilities, we observed using the method for fitting the fractions of material described in Section 2.5.7, we added equations based upon the equilibrium constants of the dimerization and trimerization reactions determined from 3-component reaction mixtures to the mass balance equations described in Section 2.5.7. For each reaction that produced a dimer or trimer, we calculated the residual using:

$$r = \ln \frac{a_P}{a_{R1} a_{R2} K_{eq}} \frac{N_{Tot}}{3C_0} \qquad (S2.8)$$

where *a* is the relative abundance of a given assembly species (R1/R2 are reactants and P is the product of a reaction governed by K_{eq}), N_{Tot} is the total amount of material in all assembly species calculated from the fractions of reaction material (f_i) and C_0 is the initial origami monomer concentration (assumed to be constant for each monomer). Note that a_i is a measure on the number of an assembly species, while f_i is a measure on the *amount of material* in an assembly species. If the equilibrium constant for a given reaction is set exactly to those measured in a two- or three- component reaction mixture, the residual value would be 0.

There was one equation for each of the dimerization and trimerization reactions, with redundancies since dimerization reactions occurred in multiple 3-component assembly mixtures, giving a total of 16 reactions for a total of 37 fitting equations including the mass balance equations. When we analyzed the data from the RNGB/RGNB/RGGB and NRBG/GRBN/GRBG sets of labeling schemes with the larger system of equations, we found little difference between the two sets of schemes in the calculated abundances (Fig. S2.19).

2.5.9 Kinetic simulations of tetramer assembly

The kinetics of the assembly of the four origami tiles into a tetramer superstructure was simulated using mass-action kinetics algorithms written in MATLAB. A reaction on-rate of $1\text{x}10^6\ \text{M}^{-1}\text{s}^{-1}$ was assumed for every reaction. This value for the on-rate is consistent with data from existing measurements for the on-rate of origami dimerization reactions driven by Watson-Crick hybridization[44] and is also similar to on-rates found for smaller DNA nanostructures.[161] The off-rates of individual reactions were calculated using this on-rate and the experimentally determined K_{eq}'s at each measured temperature:

$$k_{off} = k_{on} * K_{eq} \qquad (S2.9)$$

To match experimental assembly conditions, we simulated annealing conditions from 50 to 25 °C in 5 °C intervals and 15 minutes per temperature point. To check that simulations produced assembly abundances comparable to the experimentally measured ones, we determined the fraction of reaction material in each assembly species at the end of each temperature step (Figure S2.20). Experiments and simulations showed comparable assembly abundances. Slight differences could be due to the assumed constant on-rate changing the time to equilibrium, thus preventing some reactions reaching equilibrium within 15 minutes.

We used these simulations to investigate how the production rate of each species changes over time for a particular assembly temperature (Figure S2.21). In order to determine the net amount of material produced by each reaction for each temperature, we calculated the area under each production rate *vs.* time curve at that temperature using the trapezoidal method (Figure S2.22). From the integrated rates, we inferred which pathway(s) were likely to have been favored in the assembly. At all simulated temperatures less than 40 °C, the T12/T34 dimer+dimer tetramerization reaction produced more target structures than any other reaction did: this reaction produced more than 35% of the tetramers at those temperatures. The reaction with the second highest target structure production, at ~20% for temperatures less than 40 °C, was the reaction between T234 and T1. Of the trimer + monomer reactions, the T234+T1 reaction produced the most tetramer structures at all simulated temperatures below 50 °C.

2.5.10 Supplemental Figures and Tables:

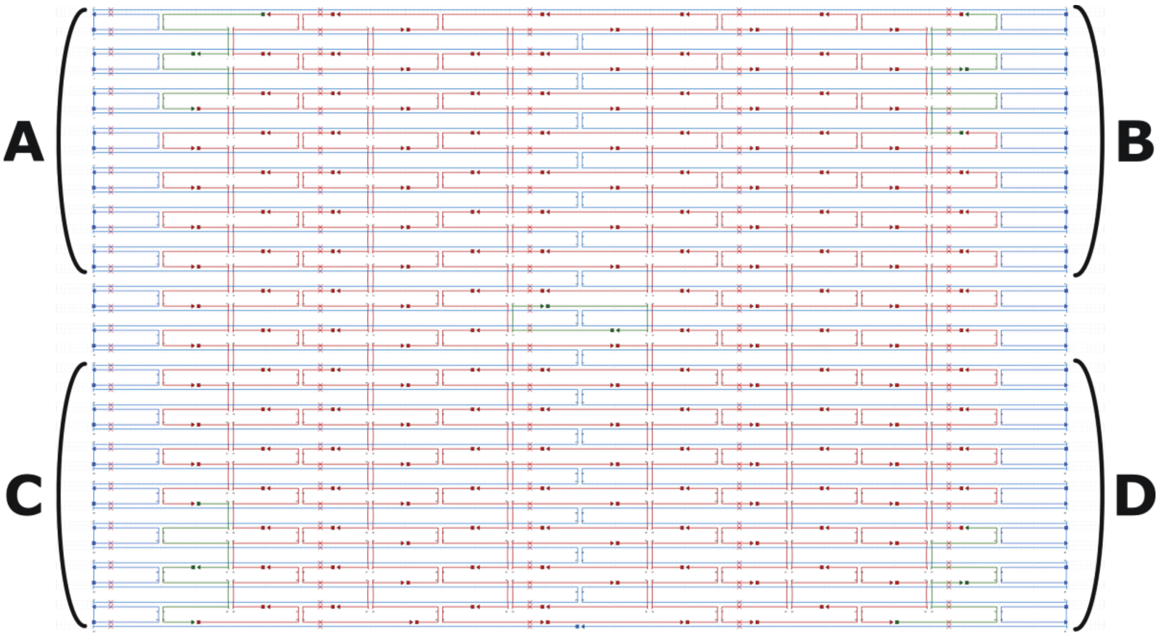


Figure S2.1: Design of the origami tile as presented in CaDNAno.[10] Edge staples = blue, surface anchoring staples = green, and structural staples = red. The architecture of the linkers and blockers on the edge of the origami tile is not shown here (see Fig. S2.3). Even though the crossover patterns are symmetric across the origami, differences in sequence, staple strand break points and hairpin locations (Fig. S2.2) create differences in structure. Due to these differences, the flexibility of the structure varies along the origami binding edges. The interfaces are labeled by their designated letter along the side of the origami.

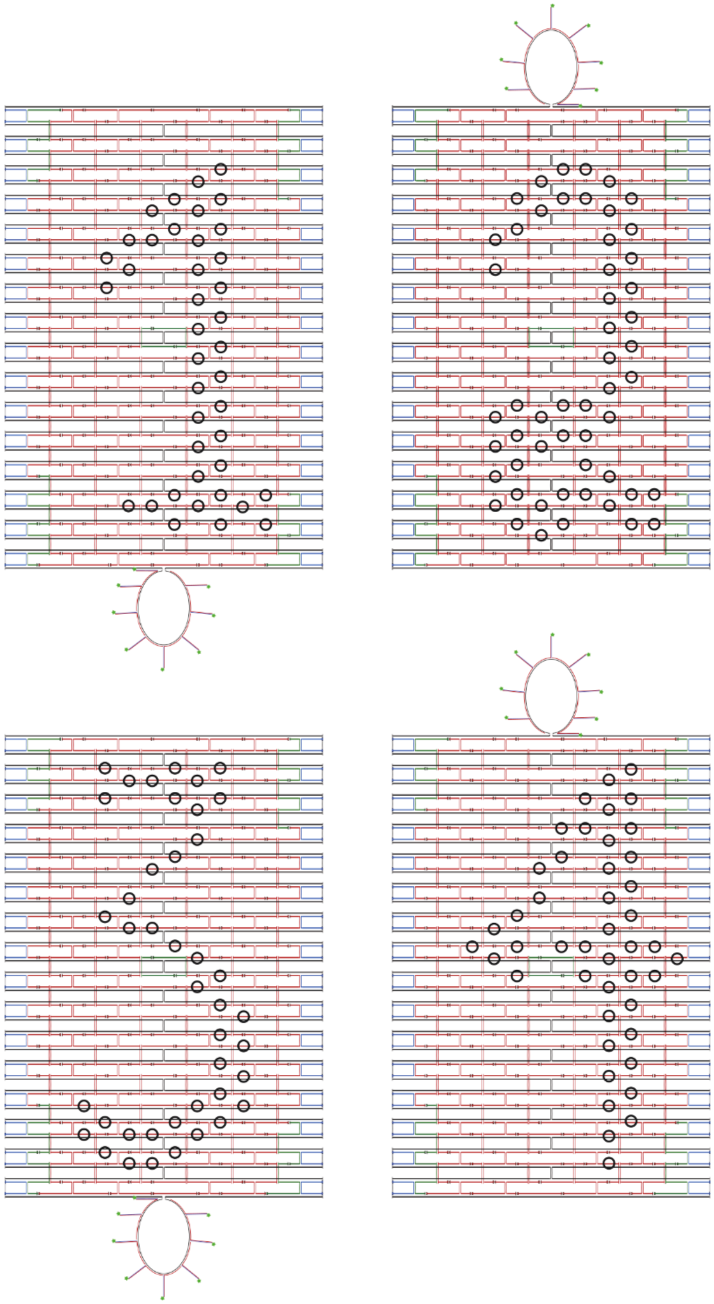


Figure S2.2: CaDNAno[10] schematics showing each origami monomer's hairpin locations. Tile 2 and Tile 4 are rotated 180° respective to Tile 1 (see Fig. S2.5). Differing hairpin locations on each origami monomer may contribute to differences in the structure of each interface.[60]

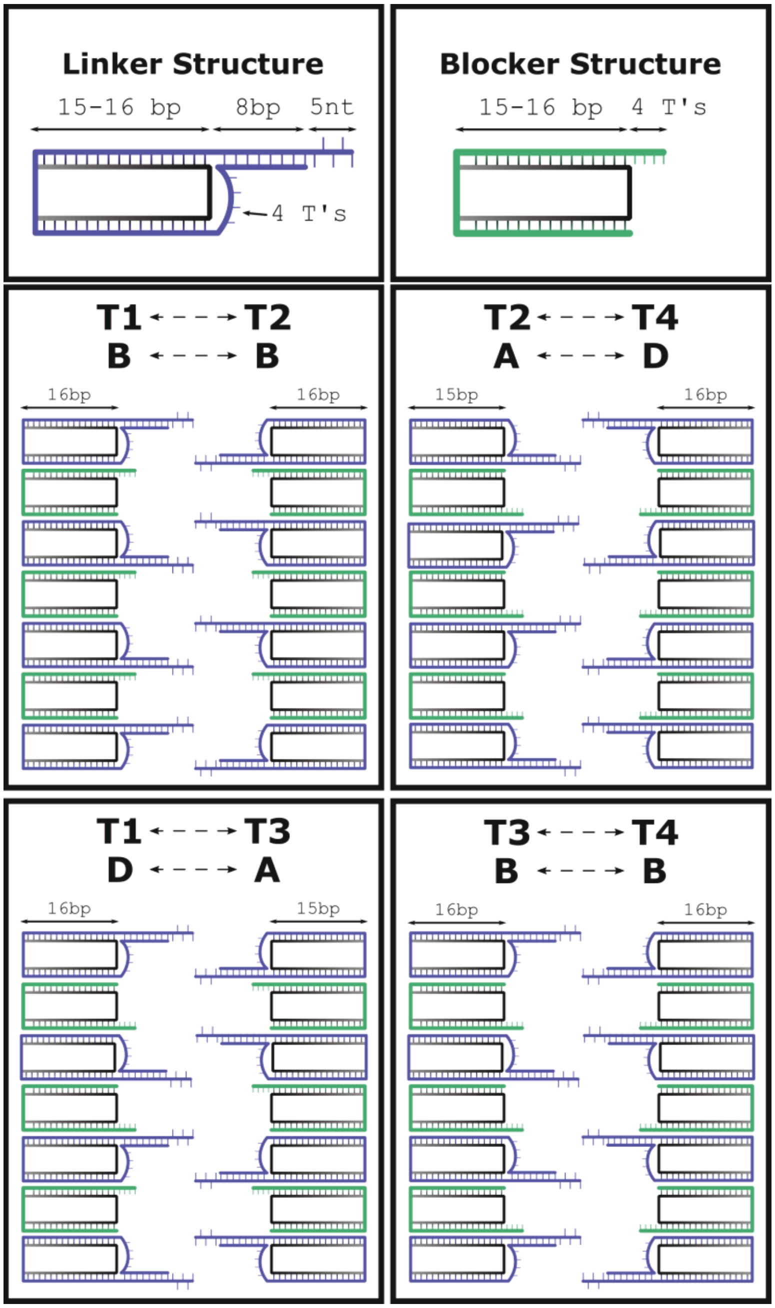


Figure S2.3: Schematics of linker and blocker edge staples are shown for each interface. Linkers are comprised of 3 parts: 30–32 bp of complementarity to the scaffold strand, 8 bp of internal complementarity along with 4 thymines of ssDNA, and 5 nt complementary to a linker strand on the opposite binding interface. As shown here, linker sticky-ends are complementary to the linker strand directly opposite each one (on a different origami). Each sticky-end is a unique sequence; these sequences are listed below and in Table S2.2. Blockers have 30–32 bp of complementarity to the scaffold and 4 thymines to block non-specific blunt-end stacking and prevent unwanted edge strand invasion from excess linker edge staples in solution. Linkers and blockers with only 30 bp instead of 32 bp of complementarity to the scaffold strand contained a deletion site in order to maintain 10.43 bp/turn helicity (see Section 2.2.1). This structural difference may also contribute to the measured differences in binding energy of each interface. The “A”, “B” and “D” denote the assigned interface type for each group of linkers/blockers.

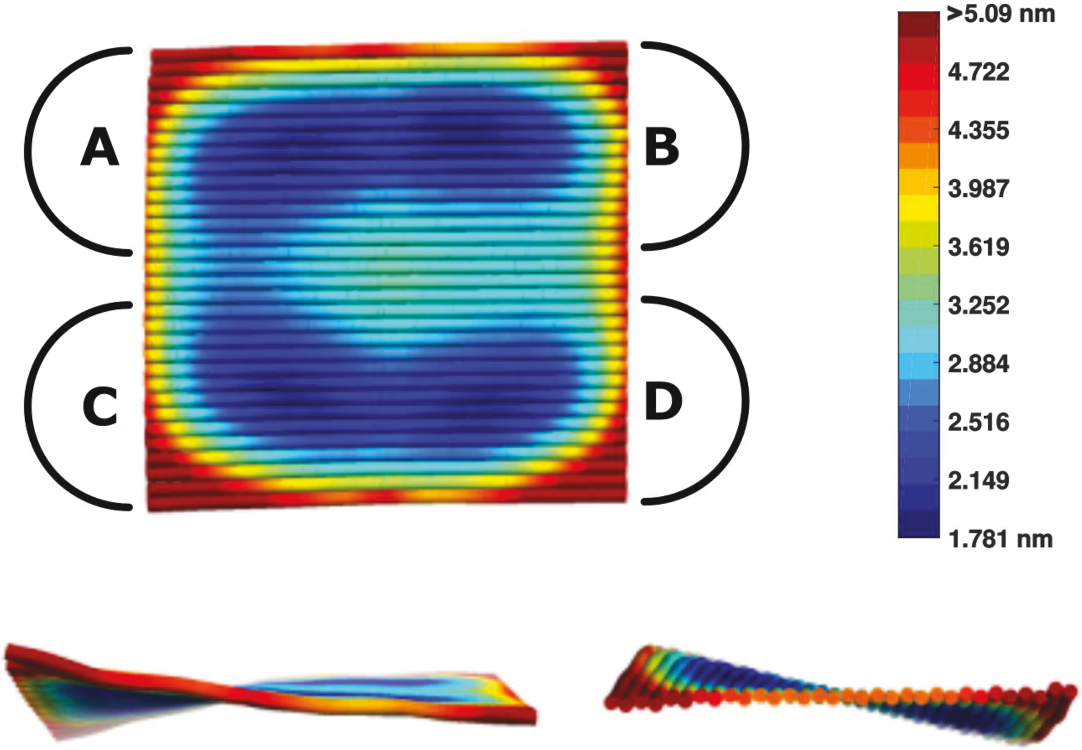


Figure S2.4: CanDo[82] predictions of the 3-dimensional structure of the origami components (without hairpins) suggest that the edges of the origami are straight, but that there is global twist parallel with the helices. Hairpins in the structure of the origami could affect the global and local twist of the origami beyond what is shown in this model.[60]

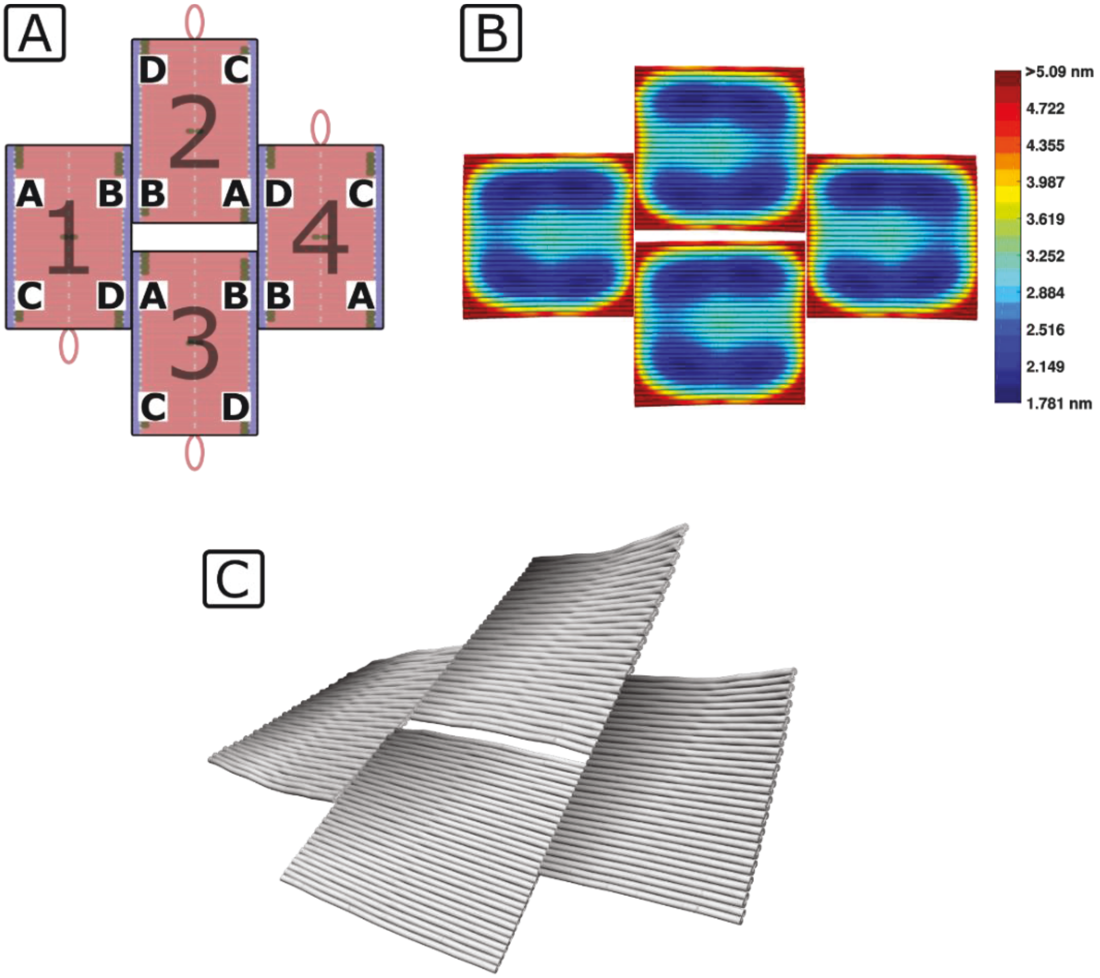


Figure S2.5: Four copies of the CanDo[82] predictions of the origami monomers (without hairpins) as shown in Fig. S2.4 aligned respective to their orientation and binding interfaces. (A) Schematic of the tetramer complex showing the relative orientation of each monomer relative to tile 1 and the corresponding interfaces involved in binding between origami components. Tile 2 and Tile 4 are rotated 180° respective to Tile 1 and Tile 3. (B) The aligned predicted structures suggest that the local flexibility, and possibly symmetry of the local structure of the origami components could play roles in determining the binding energy between components. (C) Model of the target tetramer complex showing the possible overall curvature. The models shown in (B) and (C) do not include edge or hairpin staples.

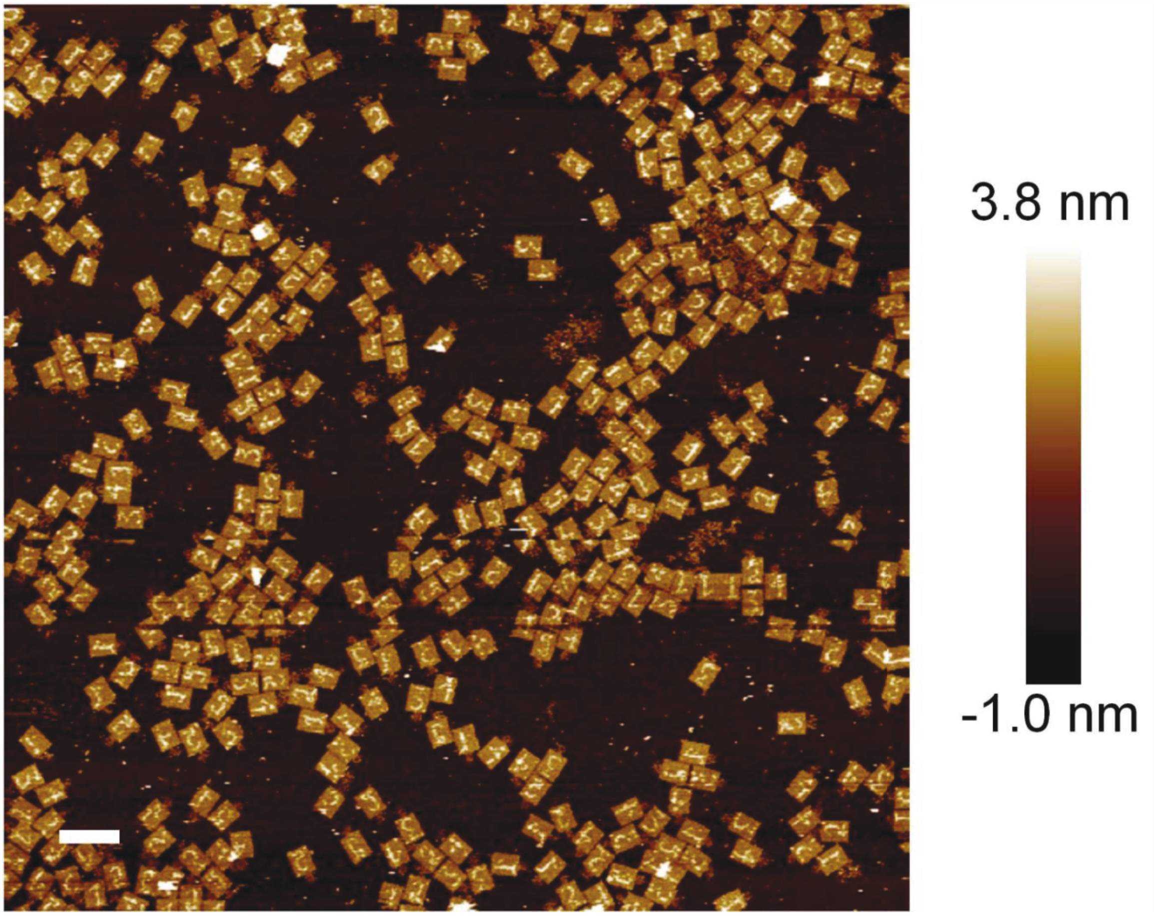


Figure S2.6: Example AFM image showing the four origami components annealed into tetramer and intermediate complexes. Most of the components are either in monomer form or within tetramers. In particular, few trimers were observed. Additionally, no incidences of hairpin patterns interchanging between origami tiles were detected, demonstrating that the annealing protocol used in this work did not commonly produce exchanges in staple strands between components. Scale bar: 200 nm.

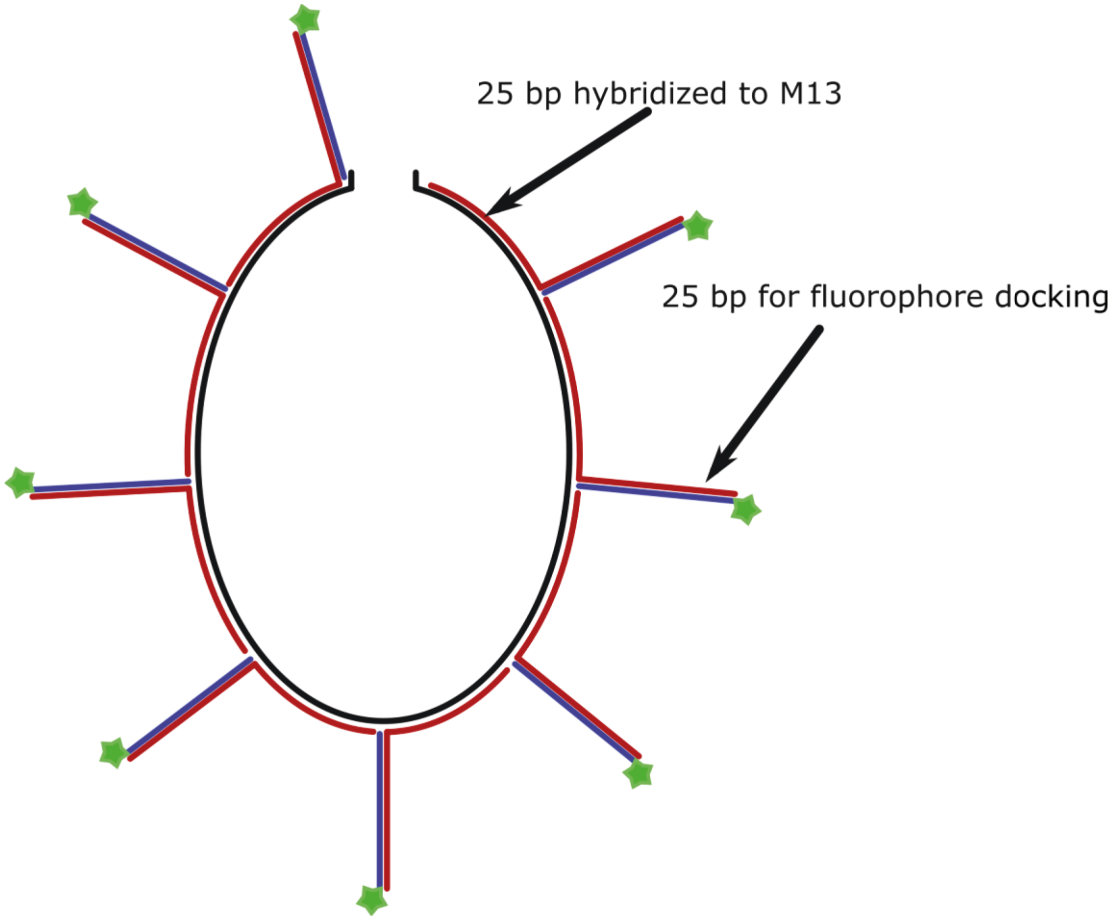


Figure S2.7: The sequence of the M13 scaffold strand not incorporated into the body of the origami structure was used as a site for fluorophore labeling. Eight docking strands (red) each with 25 bases of complementarity to the scaffold strand were hybridized to the scaffold during annealing. On fluorescently labeled origami, each docking strand consisted of this 25 base pair sequence with an additional 25 base pair sequence complementary to a fluorophore-conjugated strand (blue) which was the same for all docking strands. On origami where no fluorescent label was desired, each docking strand consisted only of the 25 base sequence complementary to its respective scaffold region: by blocking the docking sites on the M13 strand, excess fluorophore docking strands and the fluorescent strands in solution were prevented from mislabeling origami.

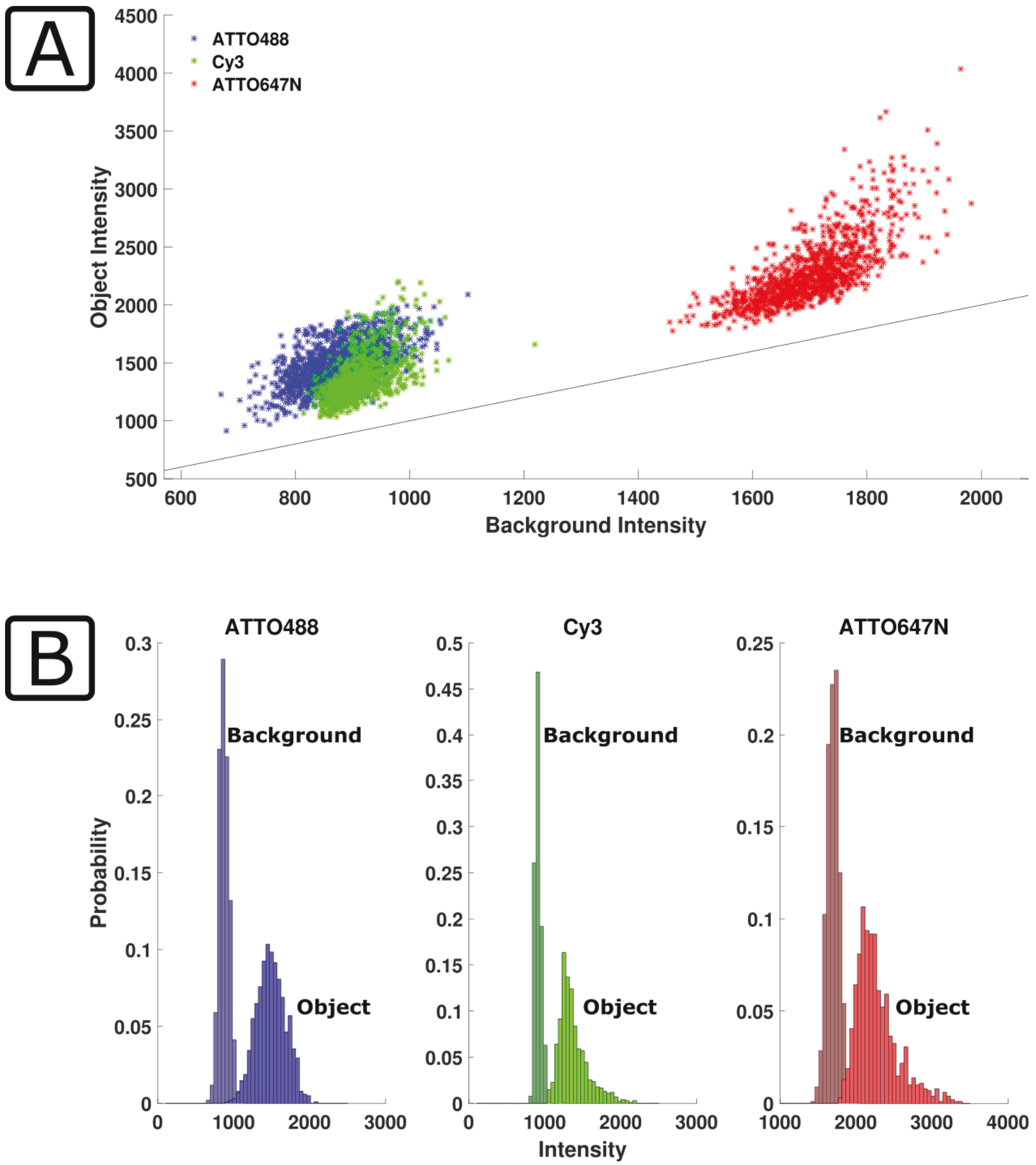


Figure S2.8: Fluorescent objects are easily detected above their local background intensity using the algorithm described in Section 2.5.3. (A) Plot of the object intensity versus each object's local background intensity (average of 3 x 3 pixel area around object). The line on the plot indicates where the local background would be equal to the object intensity. The fluorescence intensity of each object is 1.2–3x its respective local background intensity. (B) Histogram plots for each fluorophore used in the experiments. Though all objects are above a global average background, they may be at or below the local background of other objects within the same fluorophore channel.

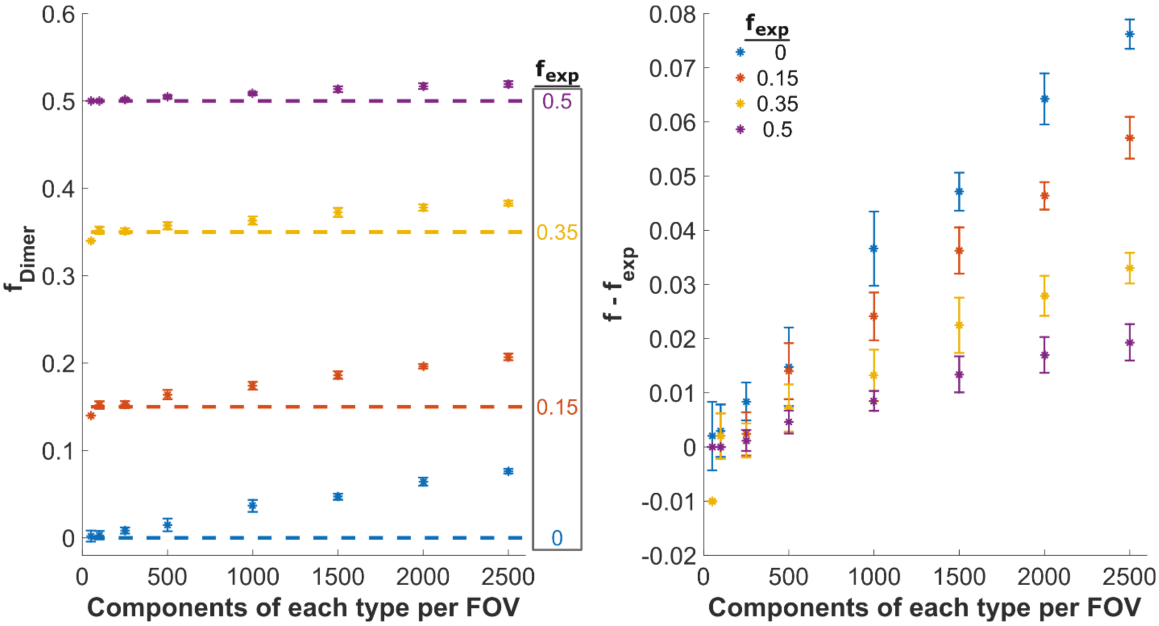


Figure S2.9: The effect of origami surface density on the simulated fraction of material in a dimer state for a two-component reaction mixture. An important potential source of error in our experiments is that fluorescent origami that are not bound to each other have a chance of landing so close to one another on the treated glass surface that they cannot be distinguished from an origami assembly. We used simulations, as described in Section 2.5.4, to determine how often such an event would occur as a function of origami surface density, assuming origami were randomly deposited on the glass surface and each component was in equal abundance. *Left*: Plots of apparent simulated fraction of material in dimer form (f_{Dimer}) for a two-component assembly reaction as a function of the number of components of each type per field of view (a 512 x 512 pixel grid). The actual expected yield (f_{exp}) of four different reaction surface depositions are depicted by four dashed lines and color coded to match the f_{exp} in the gray box. Dots with error bars show the mean and standard deviation of the fraction of material that would be measured using fluorescence co-localization as a function of component surface density on 10 simulated surfaces per dot. For high component densities, fluorescence co-localization overestimates the number of dimers because single components land near to one another on the surface even though they are not physically bound. *Right*: The difference between the measured (simulated) and actual fraction of material in a dimer state, *i.e.*, the difference between the dots and the dashed lines in the left plot, as a function of component surface density.

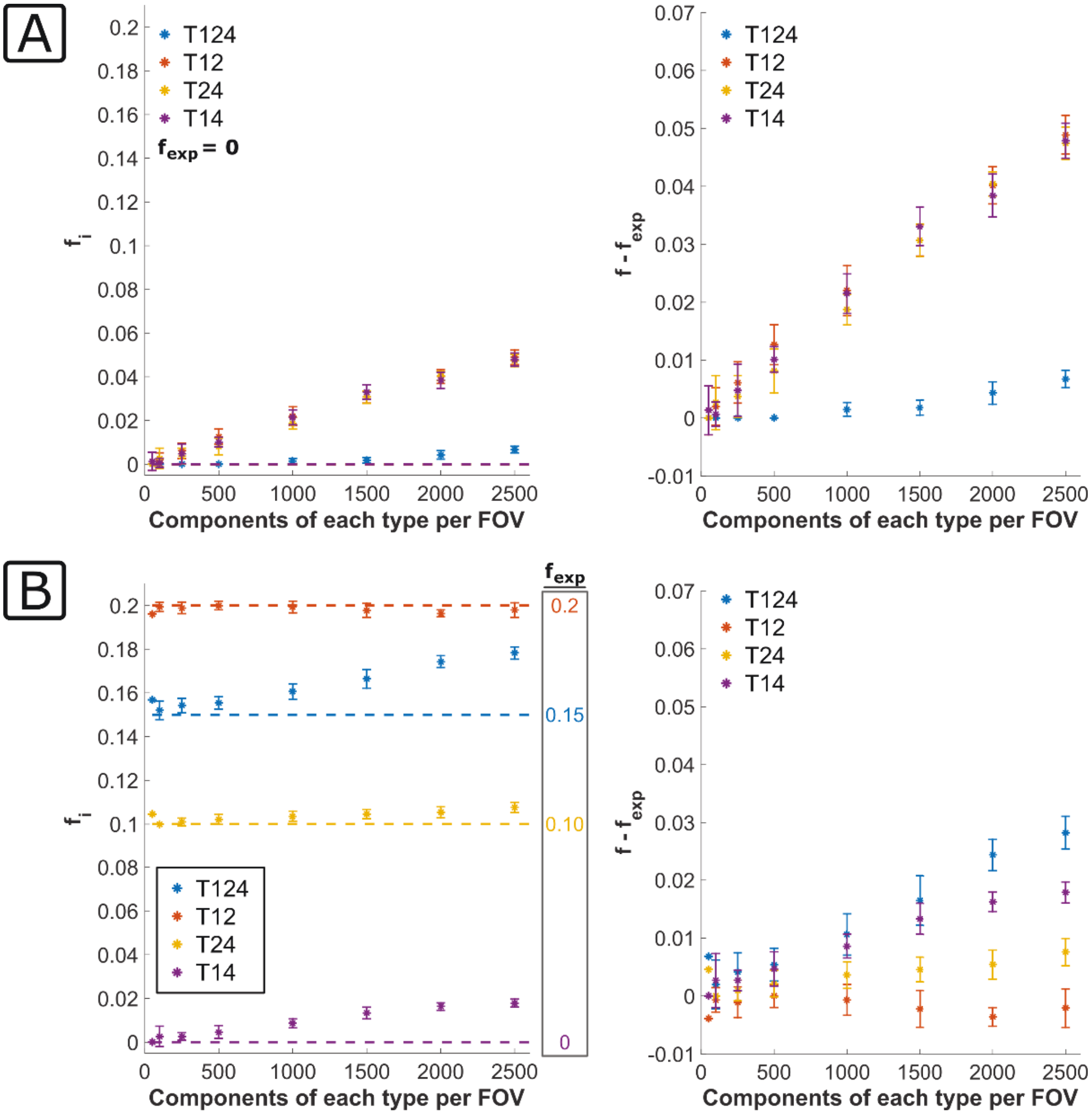


Figure S2.10: Simulation of how the chance co-localization of monomers and/or complexes in 3-component reactions could affect the measured fraction of material in each assembly state using fluorescence co-localization. (A) Simulated experiment in which 3 types of components are deposited on a surface with no expected binding between components (methods described in Section 2.5.4). The dashed line shows the expected fraction of material in dimer or trimer form (f_{exp}), which is 0 for all species. The dots (with error bars) show the average fraction of material that would be observed for each of the four species types as a function of the number of components of each type present in a single field of view. The right graph shows the error in these predictions as in Figure S2.8. (B) A simulated experiment in which 3 types of origami are allowed to react with the expected fractions of material in each of the 4 species indicated by the dashed lines and matching colored values in the gray box on the left plot. Each color corresponds to the assembly species shown in the legend (*e.g.*, T124). Left, the fraction of material that in simulation appeared to be in dimer or trimer form (dots with error bars) is plotted as a function of the number of components of each type in a field of view (*i.e.*, component surface density). The right plot shows the error for the fraction of material in each assembly type as a function of component surface density.

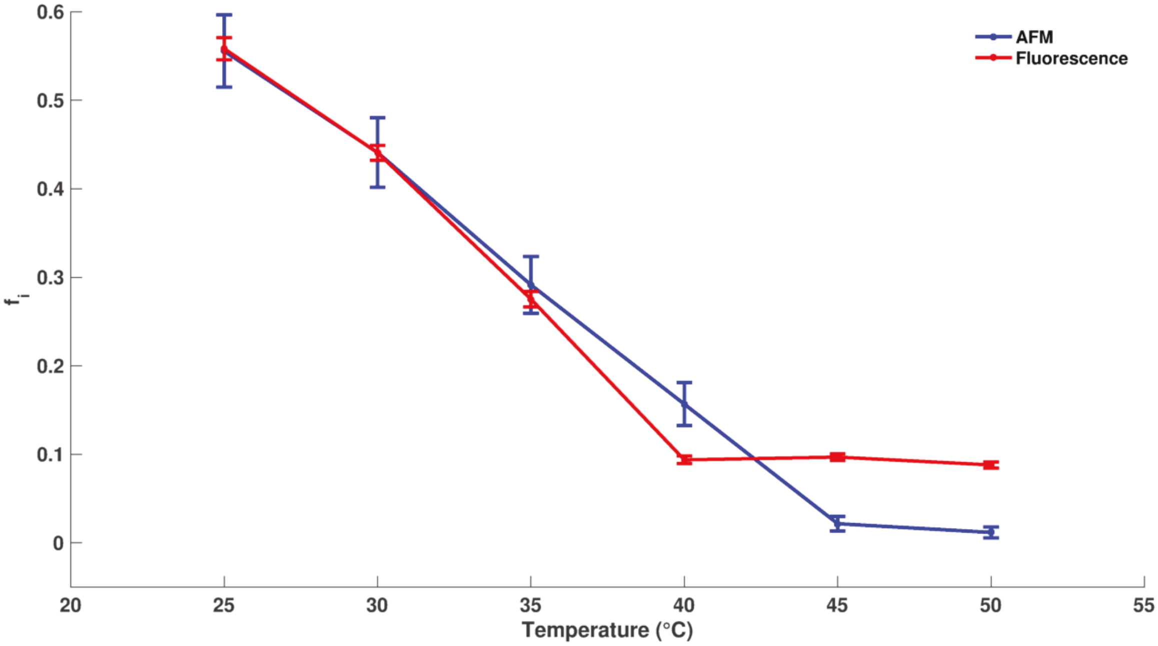


Figure S2.11: To test whether multi-color fluorescence micrographs could be used to quantitatively determine the fraction of origami components in different assembly configurations, we measured the fraction of material in each species (f_i) for a dimerization reaction between T1 and T2 using AFM (blue, no fluorophore labeling) and fluorescent microscopy (red, ATTO647N-Cy3 labeling) as a function of reaction temperature. The origami in both assays contained hairpins. Similar amounts of material in a dimer state was found in both assays at each temperature. The fraction of dimers does not reach 0 in the microscopy experiments because origami that attach to the surface randomly near to one another cannot be distinguished from assembled components (Section 2.5.4).

Table S2.1: Tabular comparison between methods of measuring the fraction of reaction material in a dimerized state for a reaction between T1 and T2. The total number of objects of a given species (n_i) is the sum across all images: 2–3 3 x 3 μm^2 or 4 x 4 μm^2 images for AFM and 6–10 87 x 87 μm^2 images for fluorescence microscopy measurements. Fractions of material (f_i) are shown as the average across all images with 95% confidence intervals determined as described in Section 2.5.5.

	AFM				Fluorescence Microscopy			
Temp.	n_{T1}	n_{T2}	n_{T12}	f_{T12}	n_{T1}	n_{T2}	n_{T12}	f_{T12}
50 °C	269	249	3	0.01±0.01	10571	13657	1171	0.09±0.01
45 °C	370	360	8	0.02±0.01	10438	11736	1189	0.10±0.01
40 °C	326	296	60	0.16±0.02	6658	9784	852	0.09±0.01
35 °C	278	232	102	0.29±0.03	3819	4166	1518	0.28±0.01
30 °C	179	153	132	0.44±0.04	4817	4694	3746	0.44±0.01
25 °C	141	115	160	0.55±0.04	1928	1792	2351	0.55±0.01

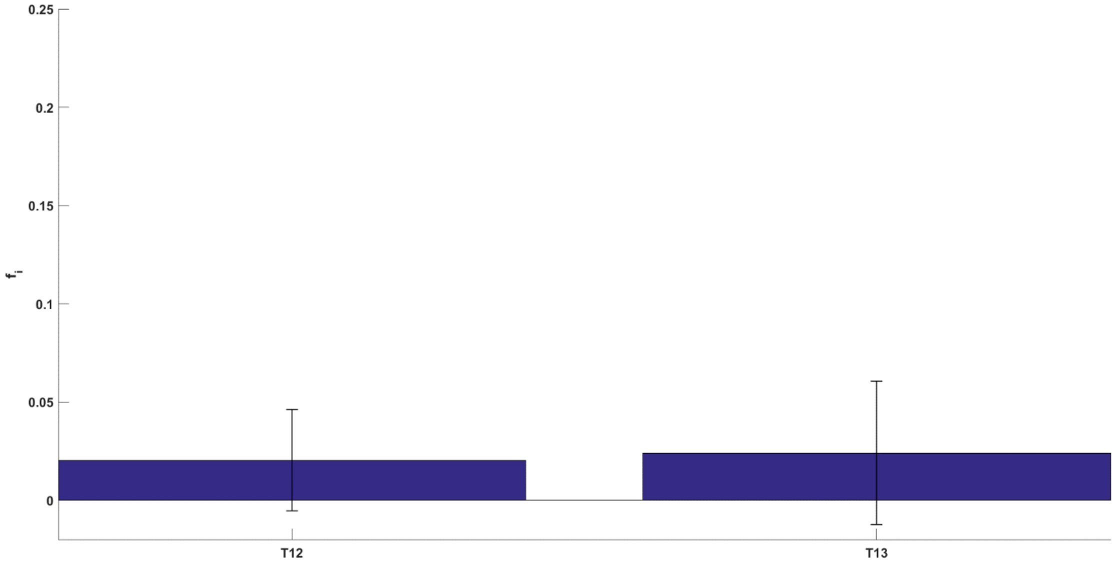


Figure S2.12: Minimal binding between origami was observed in the absence of linkers. Measured fraction of complexes (f_i) in reactions involving T1+T2 and T1+T3 components (measured after annealing to 25 °C). Many of the complexes observed are likely to be unbound components that landed near one another by chance (Section 2.5.4).

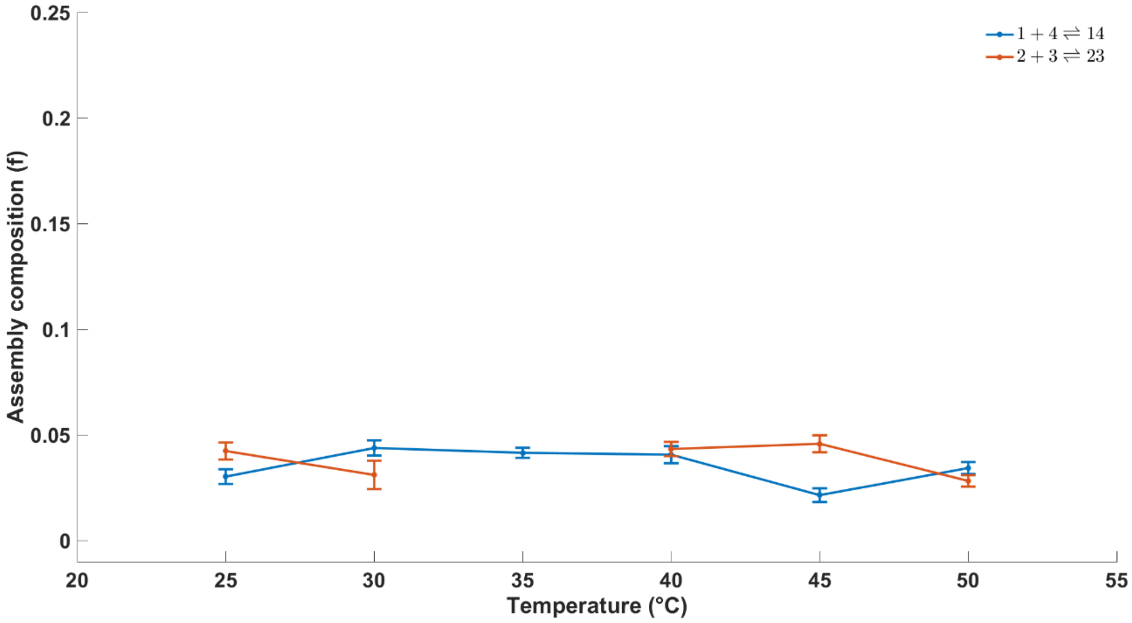


Figure S2.13: Minimal interactions were found between origami with non-complementary linking interfaces at all measured temperatures. The small number of observed assemblies is consistent with most such structures being unbound assemblies that landed near to one another by chance (Section 2.5.4). For the T2+T3 reaction at 35 °C, measurements were unable to be taken due to experimental errors.

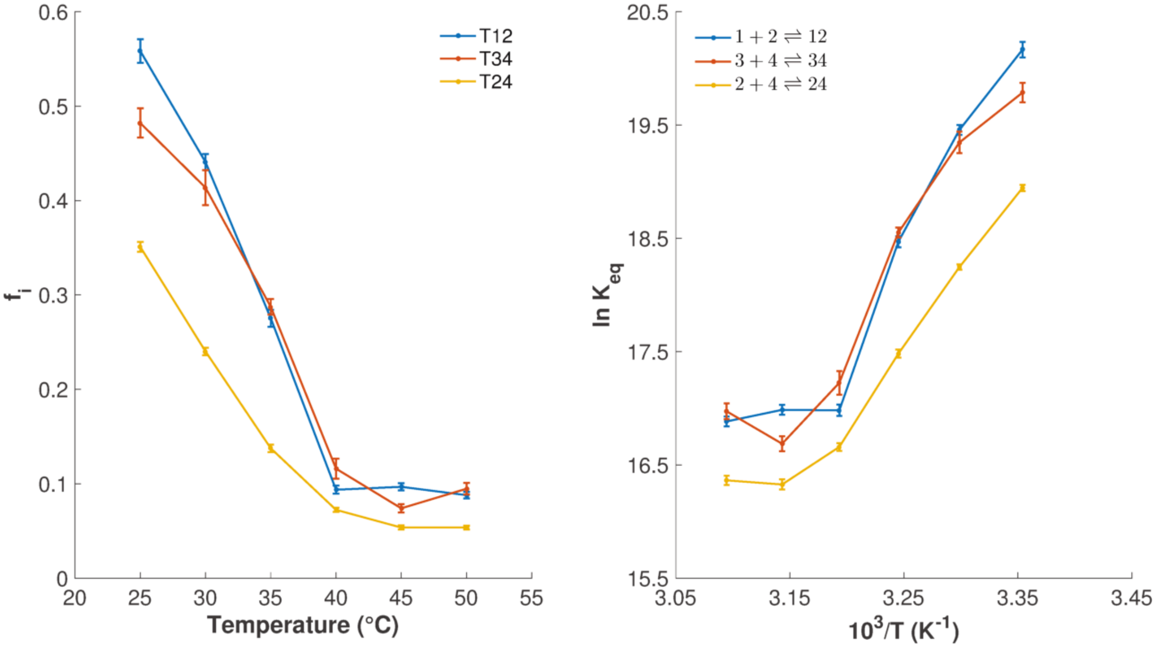
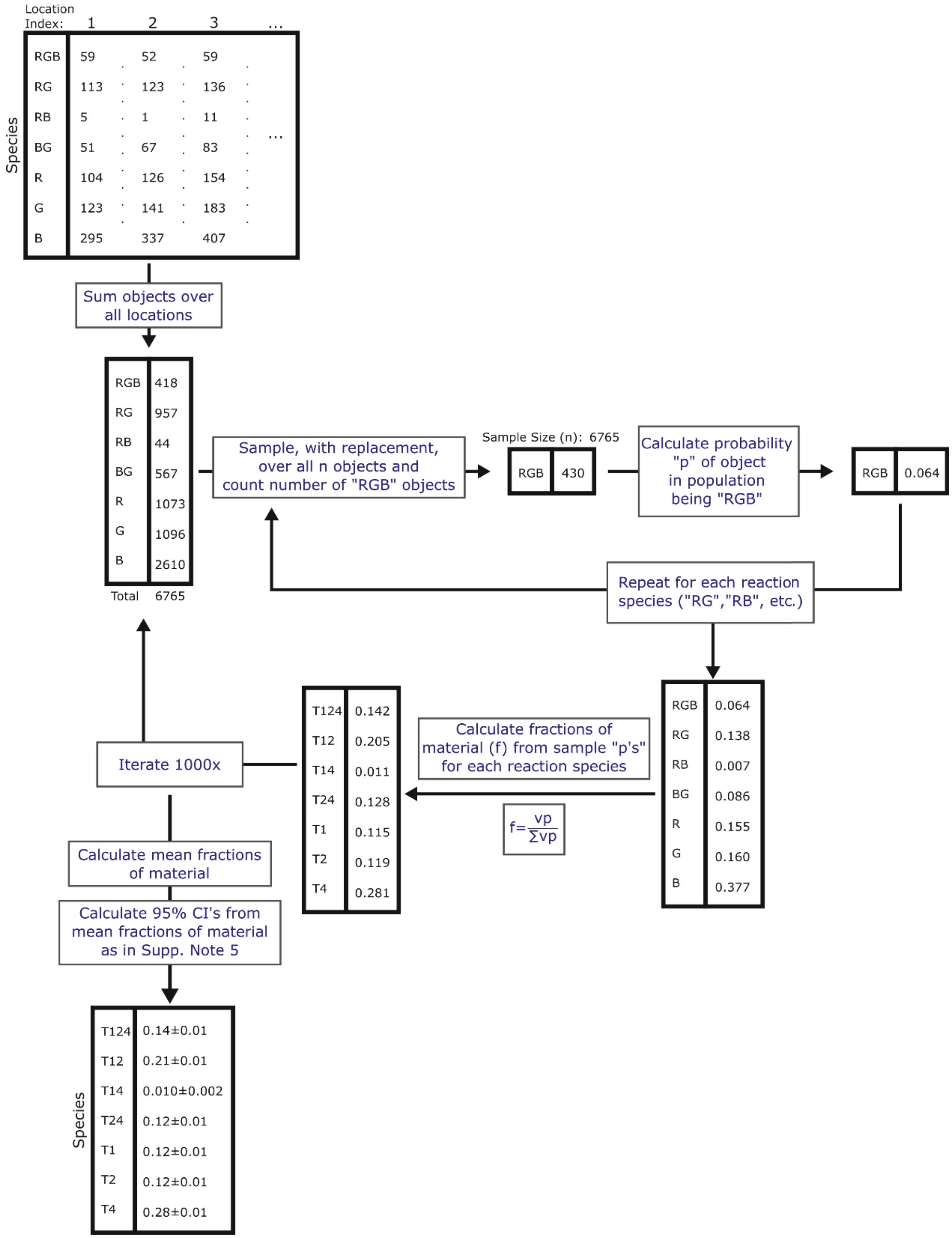


Figure S2.14: The fraction of material (f_i) of dimer species was measured for the reactions forming three different dimer products- T12, T34 and T24. The equilibrium constants for these reactions were calculated from the measured yields (Section 2.5.6).



Species	1	2	3	...
RGB	59	52	59	
RG	113	123	136	
RB	5	1	11	...
BG	51	67	83	
R	104	126	154	
G	123	141	183	
B	295	337	407	

RGB	418
RG	957
RB	44
BG	567
R	1073
G	1096
B	2610
Total	6765

RGB	430

RGB	0.064

RGB	0.064
RG	0.138
RB	0.007
BG	0.086
R	0.155
G	0.160
B	0.377

T124	0.142
T12	0.205
T14	0.011
T24	0.128
T1	0.115
T2	0.119
T4	0.281

Species	
T124	0.14±0.01
T12	0.21±0.01
T14	0.010±0.002
T24	0.12±0.01
T1	0.12±0.01
T2	0.12±0.01
T4	0.28±0.01

Figure S2.15: Bootstrapping algorithm for determining mean fractions of material in each assembly species. The 95% confidence intervals of the mean for the fractions of material in each assembly species were calculated using the equation shown as described in Section 2.5.5. The example shown is for the assembly of T124-RGB measured at 25 °C.

Table S2.2: The interaction energies between each pair of sticky-end linkers was calculated using the nearest-neighbor model.[173] The 5 bases for each sticky-end, plus one base on each side of a fully hybridized linker with its complement (underlined), were included in the calculation. Each linker interaction was designed to have an energy of around -8.5 kcal/mole, so the sum of the linkers at each interface was similar for each interface (around -33 kcal/mole). The binding energies of each interface pair cannot be assumed to be the sum of each sticky-end's hybridization energy as predicted by the nearest-neighbor model. Factors not included in the nearest-neighbor model, such as origami interface strain, linker strain and the effect of multiple linker cooperativity on binding energy, lead to differences in energy levels between the nearest-neighbor model (~ -33 kcal/mole) and the experimentally determined binding energies (~ -11 kcal/mole).

Most importantly, slight differences between interfaces with the nearest-neighbor model do not align with experimentally measured differences in the energy of origami–origami binding. The difference in binding energies between the interfaces, as measured from colocalization microscopy, is therefore unlikely to be due to differences in DNA sequence between the different reactions.

Name	Sequence	Name	Sequence	ΔG° (kcal/mole)
tile_1_1_1_5l	ATGGAC	tile_2_23_5l_beta	TCCATC	-8.23
tile_1_1_5_3l	GTCCAT	tile_2_0_24_3l	GATGGA	-8.23
tile_1_1_9_3l	CGGTAT	tile_2_0_28_3l	CATACC	-8
tile_1_23_5l_beta	CCTGTG	tile_2_0_32_5l	ACAGGG	-8.89
Total T1-T2 (Sequences)				-33.35
Total T1-T2 (B - B Interface Binding)				-11.51
tile_1_26_5l_beta	AAGTCG	tile_3_1_5l_beta	GACTTG	-8.54
tile_1_28_3l_beta	GAGAGT	tile_3_3_3l_beta	GACTCT	-8.13
tile_1_1_27_5l	AGAGTC	tile_3_5_5l_beta	ACTCTC	-8.13
tile_1_1_31_3l	CGTTGA	tile_3_7_3l_beta	CTCAAC	-8.54
Total T1-T3 (Sequences)				-33.34
Total T1-T3 (A - D Interface Binding)				-10.95
tile_2_7_3l_beta	GCAGTA	tile_4_0_2_3l	GTACTG	-8.28
tile_2_1_23_5l	CTAACG	tile_4_0_6_5l	GTTAGG	-8.07
tile_2_1_27_5l	CTGAGA	tile_4_0_10_5l	CTCAGA	-8.08
tile_2_1_31_3l	CTTGTC	tile_4_26_3l_beta	CGACAA	-8.54
Total T2-T4 (Sequences)				-32.97
Total T2-T4 (A - D Interface Binding)				-10.69
tile_3_17_5l_beta	TTACCC	tile_4_23_5l_beta	GGTAAC	-8.02
tile_3_19_3l_beta	CCCTGT	tile_4_21_3l_beta	CACAGG	-8.89
tile_3_21_3l_beta	ACTGAG	tile_4_0_28_3l	ACTCAG	-8.02
tile_3_23_5l_beta	AACTGG	tile_4_0_32_5l	CAGTTG	-8.14
Total T3-T4 (Sequences)				-33.07
Total T3-T4 (B - B Interface Binding)				-11.67

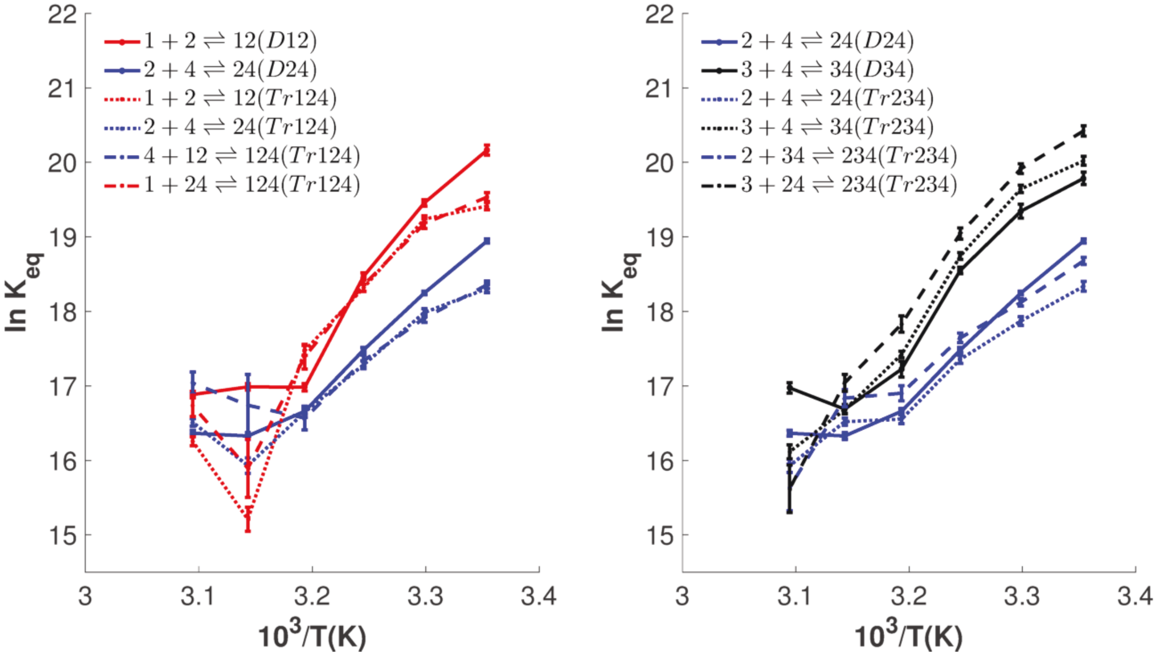


Figure S2.16: To investigate whether the equilibrium state of a reaction was affected by the presence of secondary reactions involving other reaction sites on the origami components, we compared the equilibrium constants for dimerization reactions in a two-component mixture to the equilibrium constants of the same reactions in a three-component mixture where all three components could assemble into a complex. Equilibrium constants for the possible reactions between T1-T2-T4 (left) and for the possible reactions between T2-T3-T4 (right) both suggested that binding of an origami at an interface was largely independent of binding at other interfaces on the same origami. Lines are color coded by the types of origami interfaces involved in a given assembly reaction. Solid lines: dimerization reaction in 2-component mixture; Dotted lines: dimerization reaction in a 3-component mixture; Dashed lines: trimerization reaction in a 3-component mixture.

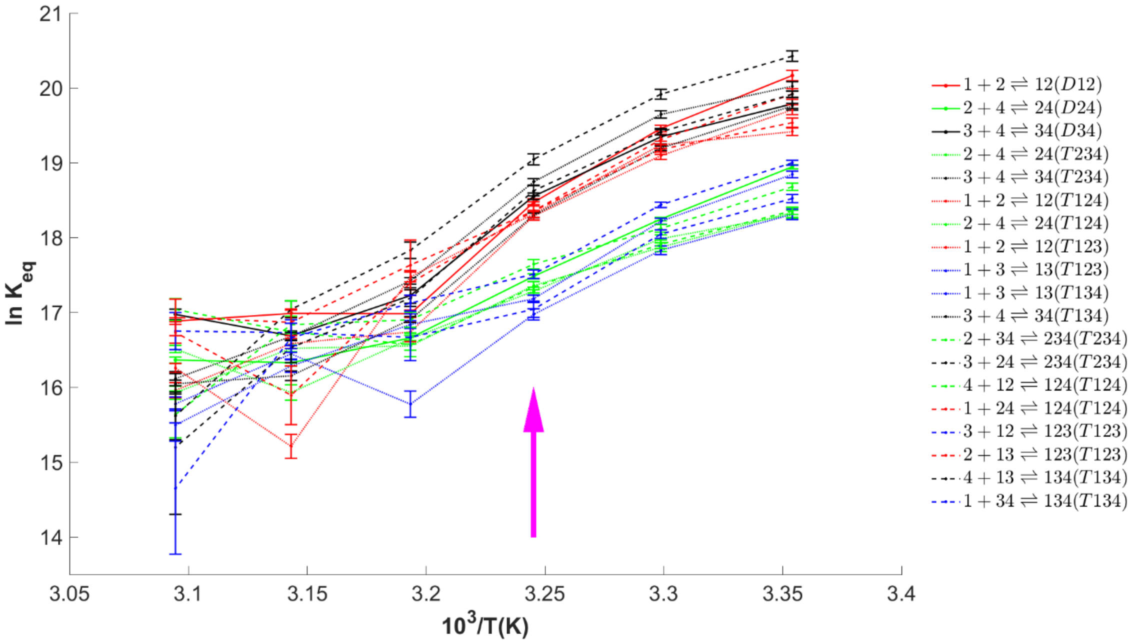


Figure S2.17: Equilibrium constants as a function of temperature, reaction, and number of components in a reaction mixture (components in each reaction solution are shown in parentheses for each reaction, D = 2 component mixture, T = 3 component mixture). At temperatures at or below 35 °C (magenta arrow), the equilibrium constants diverged into two groups depending on the pair of reacting interfaces, independent of the presence of non-reacting species. Reactions involving two “B” interfaces (red and black lines) have higher equilibrium constants than reactions involving one “A” and one “D” interface (blue and green lines). Solid lines: dimerization reaction in 2-component mixture; Dotted lines: dimerization reactions in a 3-component mixture; Dashed lines: trimerization reaction in a 3-component mixture.

Table S2.3: Three fluorescence labeling schemes for the four component origami tiles and the sets of potential assemblies and intermediates corresponding to each possible color combination in those schemes. Data from experiments using these labeling schemes were combined such that the fraction of the material of each species in each reaction could be calculated. Each fluorescent object (*e.g.*, R, R+B) represented a subset of the possible complexes. Dye Labeling: R=ATTO647N, G=Cy3, B=ATTO488, N=unlabeled.

Labeling				Fluorescent Object Combinations						
T1	T2	T3	T4	R	G	B	R+G	R+B	G+B	R+G+B
R	G	N	B	1; 13	2; 23	4; 34	12; 123	14; 134	24; 234	124; 1234
R	N	G	B	1; 12	3; 23	4; 24	13; 123	14; 124	34; 234	134; 1234
R	G	G	B	1	2; 3; 23	4	12; 13; 123	14	24; 34; 234	124; 134; 1234

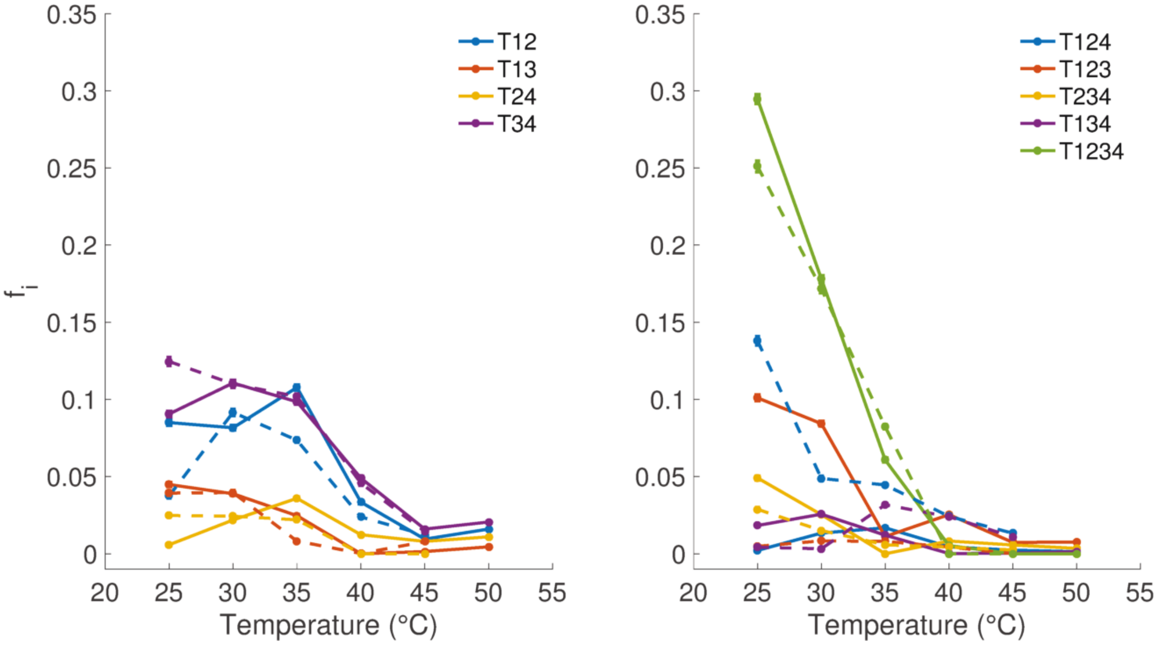


Figure S2.18: Fitting the fraction of material (fi) of each possible intermediate using two different sets of fluorescent experiments with the fitting method described in Section 2.5.7 (where only mass balance equations were used to calculate fractions of material). Plotted are the predicted fractions of material as a function of temperature for two sets of experiments with different sets of fluorescent labeling schemes. Because the fluorescent labels should not affect component binding, we would expect the results of the two sets of experiments to be the same for each component type and temperature. Results from a set of experiments using RNGB/RGNB/RGGB labels are shown by solid lines and results from a set of experiments using NRBG/GRBN/GRBG labels are shown by dashed lines. The two sets of experiments predict significantly different fractions of assemblies for T124 (right, blue) and T123 (right, red), suggesting that this fitting method may not be robust to inevitable small measurement errors.

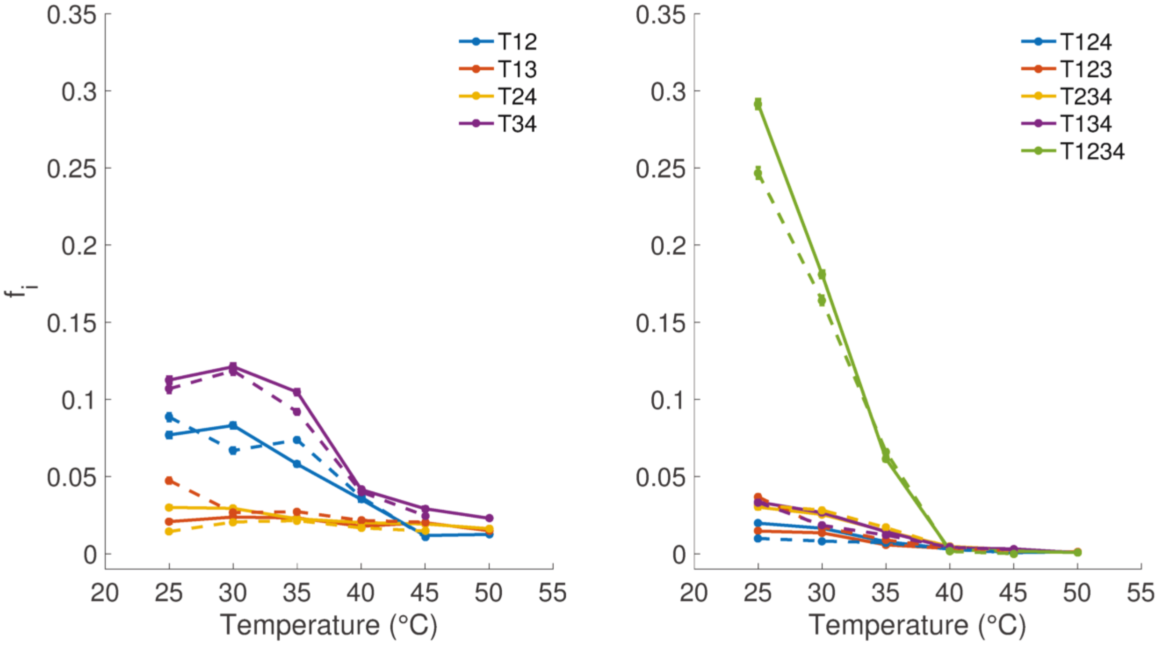


Figure S2.19: The fraction of material (f_i) in an assembly species in two sets of experiments using RNGB/RGNB/RGGB labels for components (solid lines) and NRBG/GRBN/GRBG labels for components (dashed lines) was calculated by expanding the system of equations to include the equilibrium constants of the dimerization and trimerization reactions (Section 2.5.8). The resulting fits to the data from the two sets of experiments predict very similar fractions of material of each species.

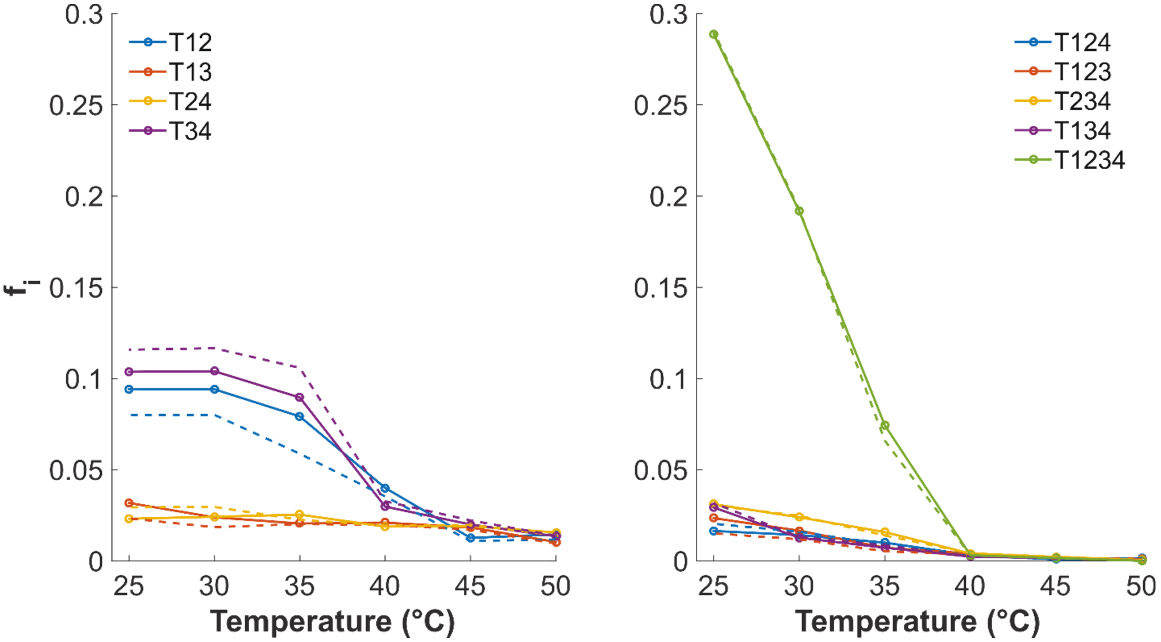


Figure S2.20: Kinetic simulations using the measured equilibrium constants, as described in Section 2.5.9, resulted in simulated equilibrium fractions of material (f_i) in each assembly complex (solid lines) similar to experimentally determined fractions (dashed lines).

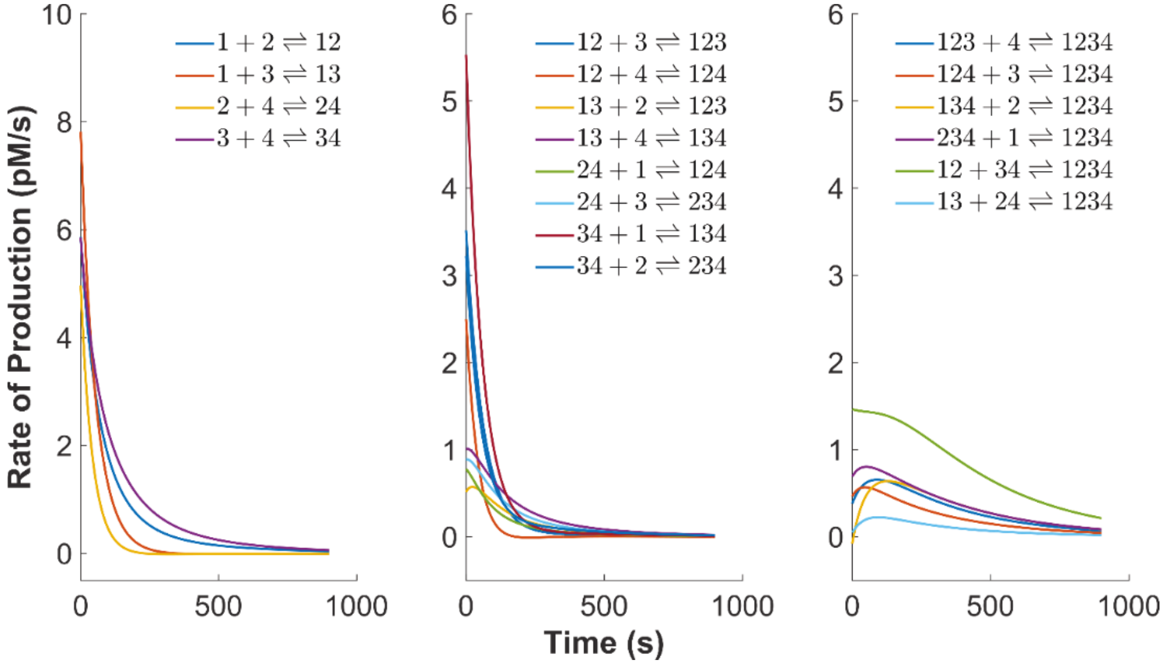


Figure S2.21: The production rate of assembly complexes as a function of time (up to 15 minutes) for assembly at 25 °C predicted using kinetic simulations (Section 2.5.9).

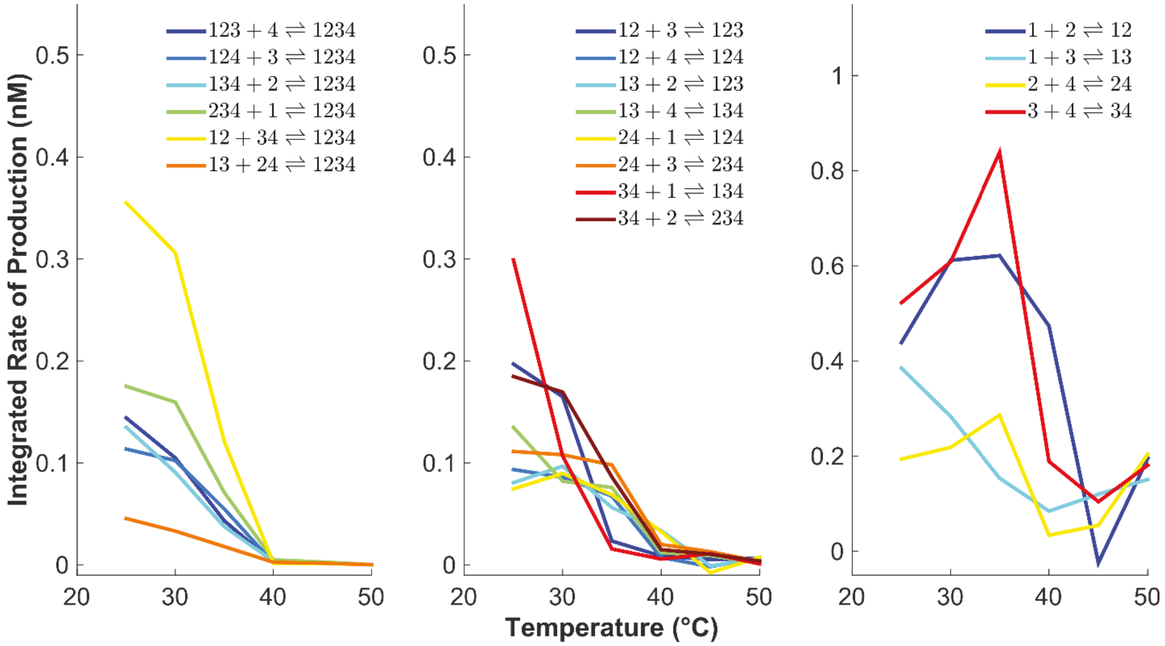


Figure S2.22: The total amount of material through each reaction during the assembly process for different temperatures as predicted by kinetic simulations of the assembly process (Section 2.5.9). These plots suggest that the dominant reaction pathways for the tetramer assembly change with temperature.

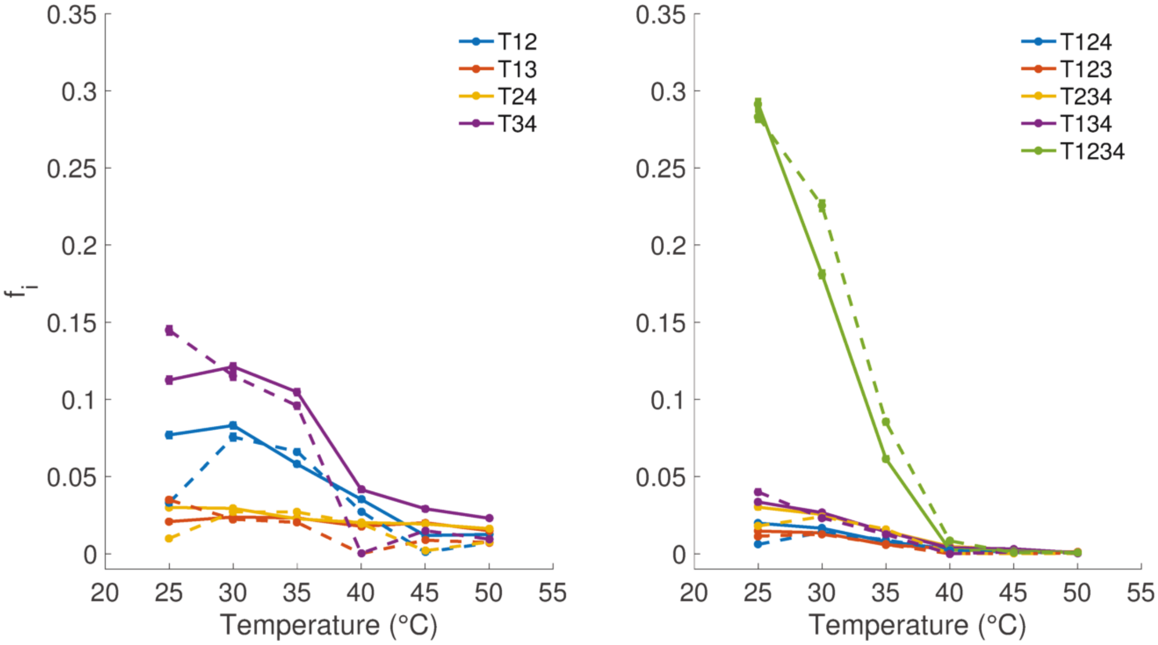


Figure S2.23: The annealing schedule for origami assembly permits at or near equilibrium analysis for the assembly of origami components. The fraction of reaction material in each assembly species (f_i) was measured during cooling (solid lines) and again during re-heating (dashed lines) using a temperature change of 15 °C/min (Section 2.5.2) for a four-component reaction mixture using the RNGB/RGNB/RGGB labeling schemes. In between annealing and re-heating, the reaction solutions incubated at 25 °C overnight. The close match between annealing and re-heating for most temperatures indicate that the annealing schedule enables each reaction to be at or near equilibrium when measurements were taken. The equilibration of the reactions on the annealing schedule used are also compatible with recently measured rates of reaction and rapid equilibration for the hybridization of two origami components.[44]

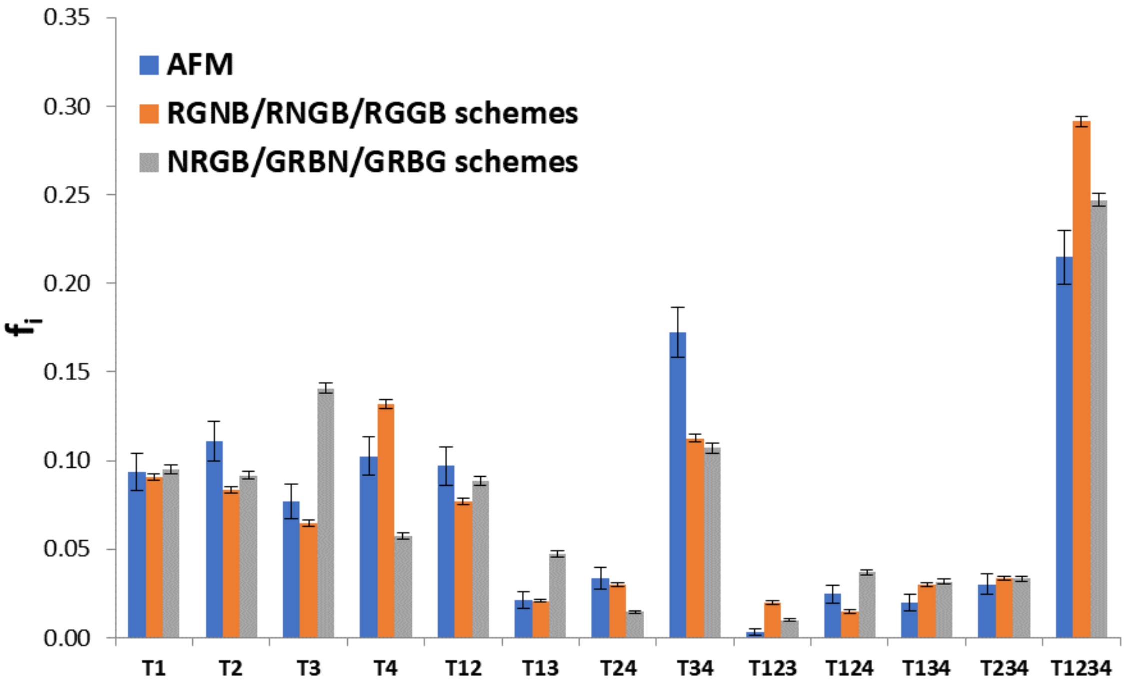


Figure S2.24: Comparison of the fraction of material (f_i) in each of the possible assembly configurations measured by AFM (blue bars) or fluorescent microscopy using two different component labeling schemes (red and green bars). To test whether the fraction of material for different species measured using fluorescence microscopy (analyzed as described in Section 2.5.8) accurately represented the actual fractions of material of the different species, we measured the relative abundance of the components labeled using AFM of the T1234-RNGB labeling scheme after annealing to 25 °C. Seven 3 x 3 micron AFM images were acquired and the number of each assembly species was counted in each image (1670 total assembly species). Independent of the method used, the major dimer species observed were T12 and T34, the tetramer target complex had a relative abundance of 25–30% and very few trimer species were observed. While AFM alone could in principle be used to could be used to analyze these systems, the amount of time required to acquire a statistically significant number of objects/images, the time required to manually count origami components and the dependency on the consistency of imaging conditions (*e.g.*, tip sharpness) to distinguish individual origami components decreases the feasibility of using AFM to analyze reaction systems of this scale and larger.

Table S2.4: Comparison of AFM and fluorescent microscopy for the determination of the fraction of reaction material in each assembly state for the formation of the tetramer complex measured at 25 °C. The average fractions of material of each species (f_i) plus/minus their 95% confidence intervals are shown. Total species counts represent the sum of each complex or monomer (*e.g.*, $n_{T1}+n_{T2}+\ldots$) for AFM measurements. For fluorescence microscopy measurements, the total species count represents the total number of co-localized objects in all three labeling schemes (*e.g.*, $n_{RGB}+n_{RG}+n_{RB}+\ldots$ in T1234-RGGB, RGNB, and RNGB labeling schemes).

	RGGB (AFM)		RGGB (Fluoro)	GRBG (Fluoro)
Species	f_i	n_i	f_i	f_i
T1	0.09±0.01	253	0.090±0.002	0.095±0.002
T2	0.11±0.01	296	0.083±0.002	0.092±0.002
T3	0.08±0.01	205	0.065±0.002	0.141±0.003
T4	0.10±0.01	275	0.131±0.002	0.057±0.002
T12	0.10±0.01	129	0.077±0.002	0.089±0.002
T13	0.02±0.01	28	0.021±0.001	0.047±0.002
T24	0.03±0.01	45	0.030±0.001	0.014±0.001
T34	0.17±0.01	228	0.113±0.002	0.107±0.003
T123	0.003±0.002	3	0.020±0.001	0.010±0.001
T124	0.03±0.01	21	0.015±0.001	0.037±0.002
T134	0.02±0.01	17	0.030±0.001	0.032±0.001
T234	0.03±0.01	26	0.034±0.001	0.033±0.001
T1234	0.22±0.02	144	0.291±0.003	0.247±0.004
Total Species Count		1,670	28,689	17,794

Table S2.5: List of each assembly reaction studied (by target complex) with the corresponding number of replicate experiments performed.

Target Complex	Replicates
T12	1
T24	3
T34	1
T123	1
T124	2
T234	2
T134	2
T1234	2

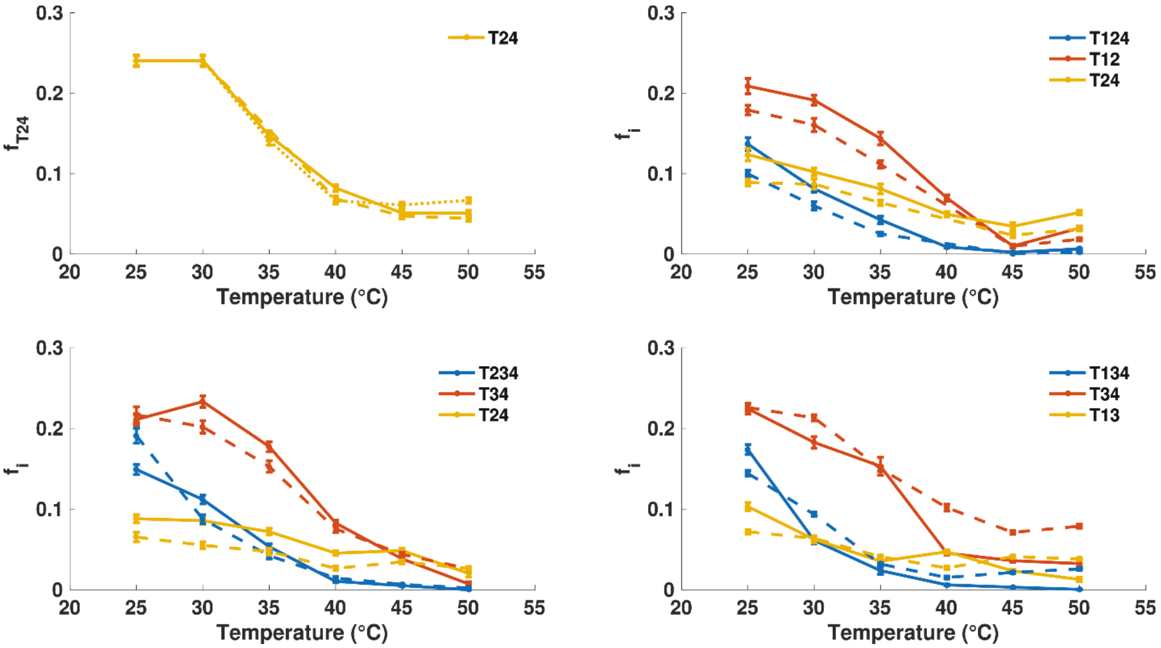


Figure S2.25: Comparison of replicate experiments for reaction mixtures with target structures T24, T124, T234 and T134. The variation between replicates appears to increase with the number of components. Error bars were calculated as described in Section 2.5.5.

Table S2.6: Filter cubes used for fluorescent imaging.

Fluorophore	Company	Set Name	Ex	Dichroic	Em
ATTO488	Semrock	FITC-2024B-OMF	FF02-485/20-25	FF506-Di03-25x36	FF01-524/24-25
ATTO647N	Semrock	LF635-A-OMF	FF01-640/14-25	Di01-R635-25x36	FF01-676/29-25
Cy3	Semrock	Custom	FF01-562/40-25	FF593-Di03-25x36	FF01-625/20-25

2.5.11 Sequences of DNA strands:

1: Fluorophore labeling strands

Unused_m13mp18_58_25v1

TAACCACCACACCCGCCGCGCTTAATTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_59_25v1

TGGCAAGTGTAGCGGTCACGCTGCGTTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_60_25v1

AAGCGAAAGGAGCGGGCGCTAGGGCTTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_61_25v1

CGAACGTGGCGAGAAAGGAAGGGAATTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_62_25v1

GATTTAGAGCTTGACGGGGAAAGCCTTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_63_25v1

TAAATCGGAACCCTAAAGGGAGCCCTTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_64_25v1

TTTTGGGGTCGAGGTGCCGTAAAGCTTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_65_25v1

TACGTGAACCATCACCCAAATCAAGTTGGACCTGACGCCTCCAGTGCCTACC

Unused_m13mp18_58_null	TAACCACCACACCCGCCGCGCTTAA
Unused_m13mp18_59_null	TGGCAAGTGTAGCGGTCACGCTGCG
Unused_m13mp18_60_null	AAGCGAAAGGAGCGGGCGCTAGGGC
Unused_m13mp18_61_null	CGAACGTGGCGAGAAAGGAAGGGAA
Unused_m13mp18_62_null	GATTTAGAGCTTGACGGGGAAAGCC
Unused_m13mp18_63_null	TAAATCGGAACCCTAAAGGGAGCCC
Unused_m13mp18_64_null	TTTTGGGGTCGAGGTGCCGTAAAGC
Unused_m13mp18_65_null	TACGTGAACCATCACCCAAATCAAG

M13Labeling_Cy3_25v1 /5Cy3/AAGGTAGGCACTGGAGGCGTCAGGTCC

M13Labeling_ATTO647N_25v1 /5ATTO647NN/AAGGTAGGCACTGGAGGCGTCAGGTCC

M13Labeling_ATTO488_25v1 /5ATTO488NN/AAGGTAGGCACTGGAGGCGTCAGGTCC

2: Biotin anchor strands

Biotin Anchor Strand:

BioAnchorv1 5' (bio) /5Biosg/TAGGTCCAAG

Added Strands: To All Tiles

2[47]_0[48]	AGT AAA TGC ATT TTC AGG GAT AGC CAC CGT AC
BAnchs29_208_26_208	GCT CCG CTT AAC GCT CAT GAA ATG GAT TAT TTA TCG CCA TTA AAA ATA GGT GAG G
BAnchs31_192_29_207	ACA GGA GGC CGA TTA AAG GGA TTA GTG TTT TTA TAA TCC AGG AAA TTG CGG AGC CTT GGA CCT A
BAnchs3_208_0_208	GTC GTC CTT ACC ACC CTC CGC CAG CAT TGA CAG GGG TCA GTG CCT TGC CTG CCT A
BAnchs16_127_15_111	TGG CTC GTT CGC CAT ATT TAA TTG CTG AAT ATC GGA TGG
BAnchs15_112_16_128	CTT AGA GCT TAA CAA CGC CAA CAT TTG AGA ATT TCG AGC CAC TTG GAC CTA
BAnchs25_32_28_32	CTT CTG GTC GTT GTA AAA CGA CGG ACT CTA GAG GAT CCC CCA CTG CCC TTG ACC TGC CTT GGA CCT A
BAnchs28_31_31_31	GCA GGT CTT GCT TTC CAA TGA ATC GGC CAA CGC GGT GGT TCC GAA ATC GCC GAG ATA
BAnchs2_31_5_31	GCG TGG CTT GTA TGG GAG GAG TGA GAA TAG AAA AAC AAC CAT CGC CCA CTG CAG GGA
BAnchs0_47_2_32	TCA GGA GGT TTA GTA CCG CCA CCC TCA GAG CCA CCA CCC TAA TTT TCT TTG CCA CGC CTT GGA CCT A
BAnchs6_207_3_207	TTA CCA GCC GGA AAC GTC ACC AAG CGA CAG AAT CAA GCA GAG CCT TGG ACG ACC TTG GAC CTA

Removed Strands: From All Tiles, Non-Hairpin strands

3[32]_5[31]

4[207]_2[208]

0[41]_1[31]

2[47]_0[42]

30[207]_28[208]

27[32]_29[31]

31[192]_30[208]

2[207]_0[208]

29[32]_31[31]

1[32]_3[31]

28[207]_26[208]

Removed Strands: T1/T2 (non-hairpin strands); T3/T4 (hairpin strands)

15[128]_16[112]

16[111]_15[127]

3: Structural staple strands

Hairpin Sequence (see Rothemund, 2006):[7] TC CTC TTT TGA GGA ACA AGT TTT CTT GT

Non-Hairpin staples:

23[160]_25[159]	GGA GCG GAC TCG TAT TAA ATC CTT TGG CAA
3[32]_5[31]	AAT AGA AAA ACA ACC ATC GCC CAC TGC AGG GA
29[192]_31[191]	CAT TGC AAA GTG AGG CCA CCG AGT GGA GCT AA
4[207]_2[208]	AAT CAA GCA GAG CCA CCA CCC TCC GCC AGC
10[207]_8[208]	TGT TTA AAT AAG AGC AAG AAA CCT TTT TAA
6[207]_4[208]	TTA CCA GCC GGA AAC GTC ACC AAG CGA CAG
19[160]_21[159]	TTA ATT TTA TTT GAA TTA CCT TTA CAA TAA
0[175]_1[159]	TAT TAA GAG GCT GAG ACT CCT CAT ACA TGG
31[160]_30[176]	GCT TTC CTC GTT AGA ATC AGA GCG AAA AGA GT
14[143]_12[144]	GAG CAT GTA TCA TTC CAA GAA CGG GGG AGG TT
14[175]_12[176]	ACA AGA AAC CGC ACT CAT CGA GAA GCG AGG CG
16[63]_18[64]	ACG GTG TCC TGT TTA GCT ATA TAG CAA AAT
15[192]_17[191]	AAA AGG TAA GTA TCA TAT GCG TTA GAA ATA CC
7[160]_9[159]	AAA ATA CAA ACC GAG GAA ACG CAC GCT AAT
0[41]_1[31]	GGT TTA GTA CCG CCA CCC TCA GAG CC
26[175]_24[176]	CAC GCT GAT TGA AAG GAA TTG AGG ACA AAC AA
13[192]_15[191]	CCG TTT TTT TAT CAA CAA TAG ATA GTA CCG AC
6[95]_8[96]	AAG TTT CCT GTC GAA ATC CGC GAC GCA GAC GG
15[32]_17[31]	GGA AGC AAA TAT AAC AGT TGA TTC ACC ATT AG
18[207]_16[208]	GAG ACT AGA TAA ATA AGG CGT TAG AAA AAG
13[160]_15[159]	ACC AAG TAA ATA ATA TCC CAT CCC ATT TTC
25[32]_27[31]	CTT CTG GTC GTT GTA AAA CGA CGG ACT CTA GA
6[175]_4[176]	CAA CCG ATA ATC ACC AGT AGC ACC GAC TGT AG
10[63]_12[64]	TGC GAT TGA TTC ATC AGT TGA GCA GAC GAC
14[207]_12[208]	CGC CTG TAT TTT CAT CGT AGG AAT ATA GAA
18[47]_16[48]	CAG AGC ATC GCA AAT GGT CAA TAA CTG GAA GT
29[160]_31[159]	CTG GTA ATC ACG CAA ATT AAC CGA TAA CGT
9[192]_11[191]	GCC CAA TAC GTC AAA AAT GAA AAT AGA GCC TA
20[63]_22[64]	ACC GTT CAT CGA TGA ACG GTA AAA CGT TAA
1[160]_3[159]	CTT TTG ATA TAT TCA CAA ACA AAA CCG GAA
9[160]_11[159]	ATC AGA GAA GAA TAA CAT AAA AAG AAT CTT
3[80]_1[79]	TTG AAA ATA GCG TAA CGA TCT AAA CCA TGT AC
4[63]_6[64]	CCG ATA GGT CAC CCT CAG CAG CAC GTA ATG
0[63]_2[64]	GGT GTA TAA GCC CAA TAG GAA CGT TTT GTC
23[80]_21[79]	GCG AGT AAA AAT ATT TAA ATT GTA TCG TAA AA
28[63]_30[64]	ACT CAC ATG CGT ATT GGG CGC CTT TGC CCC
13[32]_15[31]	ATA GCG TCC CCG AAA GAC TTC AAA ACC AGA CC

26[95]_28[96]	GGG ATG TGT GAA ATT GTT ATC CGC GCA TAA AG
19[32]_21[31]	GAA GCC TTG CCG GAG AGG GTA GCT TCA TTG CC
5[32]_7[31]	GTT AAA GGA ACG AAA GAG GCA AAA CAG CGA TT
14[95]_16[96]	CCT GAC TAT AAG AGG TCA TTT TTG AAT GCT GT
2[47]_0[42]	AGT AAA TGC ATT TTC AGG GAT AGC CAC CGT ACT CAG GA
17[80]_15[79]	GGG GCG CGA TGT TTT AAA TAT GCA TAA TTG CT
30[207]_28[208]	TAT AAT CCA GGA AAA ACG CTC ATG AAA TGG
18[95]_20[96]	ATC ATA CAA TGC AAT GCC TGA GTA AAG GCC GG
18[175]_16[176]	GGG TTA TAC TAA ATT TAA TGG TTT TAC AAA TT
3[192]_5[191]	GCC ACC CTT TTG CCT TTA GCG TCA ATT ACC AT
20[95]_22[96]	AGA CAG TCT CAA TCA TAT GTA CCC GGA AGA TT
28[143]_26[144]	CAG AGA TAT GGC ACA GAC AAT ATT GAA AAA TC
4[143]_2[144]	CCC TTA TTC GGA ACC AGA GCC ACC TAA ATC CT
22[143]_20[144]	TTT CAG GTT TTT ACA TCG GGA GAA TTT AAT GG
24[207]_22[208]	AGG ATT TGC AAT TCA TCA ATA TAT GGA AGG
18[143]_16[144]	AAA TCC AAA AAT ATA TTT TAG TTA CAA CAG TA
8[47]_6[48]	AAC GGT GTC GCG AAA CAA AGT ACA GAA GGC AC
2[95]_4[96]	TCA TAG TTC TCC AAA AAA AAG GCT TCA GCT TG
26[47]_24[48]	AGT CAC GAG CCG GAA ACC AGG CAA TGC CAG TT
13[80]_11[79]	CTC AAA TGC ATA ACC CTC GTT TAC ATT TAG GA
0[207]_1[191]	TTT CGG AAC CTA TTA TTC TGA AAG GTA ATA A
14[63]_16[64]	GCA TCA AAT TAG AGA GTA CCT TAC TAA AGT
24[143]_22[144]	ACG TTA TTG CGG AAC AAA GAA ACC GCG TAG AT
4[175]_2[176]	CGC GTT TTC TCA GAG CCG CCA CCC CAG ACG AT
17[160]_19[159]	CTT CTG ACT AAC TAT ATG TAA ATC TAT TAA
4[47]_2[48]	ACA ATG ACG GAA CAA CTA AAG GAA CAG ACG TT
18[63]_20[64]	TAA GCA ATA AAA ATT TTT AGA AGA TAT TCA
12[143]_10[144]	TTG AAG CCG CTA CAA TTT TAT CCT CAG GGA AG
17[192]_19[191]	GAC CGT GTC CTT TTT AAC CTC CGG AAA CAT AG
26[63]_28[64]	CGC CAG GTT CGT AAT CAT GGT CGT GAG CTA
30[63]_31[81]	AGC AGG CAA CAA GAG TCC ACT ATT AAA GAA CG
11[80]_9[79]	ATA CCA CAC TTT AAT CAT TGT GAA CAT TCA GT
26[207]_24[208]	CGG TCA GTC TTT AGG AGC ACT ATA CAT TTG
12[175]_10[176]	TTT TAG CGT AAC GAG CGT CTT TCC AGC AGC CT
8[175]_6[176]	CAG AAG GAT ACA TAA AGG TGG CAA GCG ACA TT
29[80]_27[79]	TTT TCT TTC TGG GGT GCC TAA TGA ATA GCT GT
7[80]_5[79]	ATA AAT TGA TTA AAC GGG TAA AAT GAA AGA CA
25[160]_27[159]	ATC AAC AGG AGC CAG CAG CAA ATT TTG AAT
28[47]_26[48]	GTT GCG CTG GGT ACC GAG CTC GAA GTT TTC CC
21[160]_23[159]	CGG ATT CGA GAA ATA AAG AAA TTA CCA GAA
22[47]_20[48]	CGC ATT AAT GGA GCA AAC AAG AGA TAG CTG AT
19[192]_21[191]	CGA TAG CTA GAA AAC AAA ATT AAT ACC AAG TT
0[143]_0[112]	ATT AGG ATT AGC GGG GTT TTG CTC AGT ACC AG

21[80]_19[79]	CTA GCA TGA AAT CAC CAT CAA TAT CCC TCA TA
16[47]_14[48]	TTC ATT CCA CTC CAA CAG GTC AGG AAA GAT TA
31[144]_28[144]	GAG CAC GTT TGT AGC AAT ACT TCT AAG AAC TCA AAC TAT CCT GGC CAA
24[175]_22[176]	TTC GAC AAA TTA TCA TCA TAT TCC TAT TTG CA
7[192]_9[191]	AGA AAC GCA AGC AGA TAG CCG AAC TGA GTT AA
8[63]_10[64]	GAA AGA GCG TAA CAA AGC TGC TTT ACC TTA
12[207]_10[208]	GGC TTA TGT TAC AAA ATA AAC ACG ATT TTT
27[32]_29[31]	GGA TCC CCC ACT GCC CGC TTT CCA ATG AAT CG
27[160]_29[159]	GGC TAT TAA ATA AAA GGG ACA TTG GCC TTG
31[192]_30[208]	ACA GGA GGC CGA TTA AAG GGA TTA GTG TTT T
11[192]_13[191]	ATT TGC CAC CGG TAT TCT AAG AAC CAA GCA AG
20[207]_18[208]	AAC ATC ATA GAT TAA GAC GCT GAG GTC TGA
24[47]_22[48]	TGA GGG GAG GCC TTC CTG TAG CCA TTA AAA TT
24[63]_26[64]	GTG CAT CAG CGC CAT TCG CCA TTT GGG TAA
14[47]_12[48]	AGA GGA AGC AAT ACT GCG GAA TCG ACC AAA AT
25[192]_27[191]	CTA AAA TAT ATT AAC ACC GCC TGC CTG ATA GC
25[128]_26[112]	AAC CCT CAC ACC TTG CTG AAC CTC CTA TTA CG
6[111]_5[127]	GAC TTT TCT ACA GAG GCT TTG AAT TAT CAC
15[128]_16[112]	GCC AAC ATT TGA GAA TCG CCA TAT TTA ATT GC
24[111]_23[127]	ACC GTA ACC GTG GGA ACA AAC GGT AAC ATT
13[128]_14[112]	CTT TCC TTA GAA ACC AAT CAA TAA TAA ATC AA
8[111]_7[127]	ACG AGG CCT GCT CCA TGT TAC TAT TAA GAC
9[128]_10[112]	GAA CAA AGA CGG GAG AAT TAA CTG GTA AAT TG
14[111]_13[127]	AAA TCA GAA ACG AGA ATG ACC ATC GGC TGT
12[111]_11[127]	AGG CAT AAC ATA ACG CCA AAA GTG CTA TTT
22[111]_21[127]	AAA AAC ACG GTT GAT AAT CAG AAT ACA GTA
16[111]_15[127]	TGA ATA TCG GAT GGC TTA GAG CTT AAC AAC
3[128]_4[112]	AAA ATC ACA GCG TTT GCC ATC TTT AAT TGT AT
23[128]_24[112]	ATC ATT TTA ATT TTA AAA GTT TGA GCG GAT TG
5[128]_6[112]	CGT CAC CGT ATT CAT TAA AGG TGA GGA CTA AA
2[111]_1[127]	TTC CAC ATA CAA ACT ACA ACG CAG TCT CTG
1[128]_2[112]	AAT TTA CCC CAG AAT GGA AAG CGC CTG TAG CA
26[111]_25[127]	CCA GCT GGT GCG GGC CTC TTC GAA ATA TCA
21[128]_22[112]	ACA GTA CCT TAA CGT CAG ATG AAT AAA GCC CC
10[111]_9[127]	GGC TTG AAC ACC AGA ACG AGT AAA CAC CCT
11[128]_12[112]	TGC ACC CAT TAA ATC AAG ATT AGT GAA TTA CG
4[111]_3[127]	CGG TTT ACC AAA AGG AGC CTT TTC ATA ATC
29[128]_30[112]	CTG AGT AGT TGA TTA GTA ATA ACA CTT CAC CG
30[111]_29[127]	CCT GGC CCA ACA GCT GAT TGC CTC ACT TGC
17[128]_18[112]	ACT TTT TCT CGC AAG ACA AAG AAC GTA GTA GC
18[111]_17[127]	ATT AAC ATC AAT TCT ACT AAT AGC GAG AAA
27[128]_28[112]	AGA ATA CGG AAC CCT TCT GAC CTG ACA TAC GA

7[128]_8[112]	TCC TTA TTC AAA AGA ACT GGC ATG TAG CCG GA
20[111]_19[127]	GGT GAG AAT GTG TAG GTA AAG ATG AGT GAA
28[111]_27[127]	GCC GGA ATC ACA ATT CCA CAC AAA AGC GTA
19[128]_20[112]	TAA CCT TGC ATA AAT CAA TAT ATG TTC AAA AG
24[95]_26[96]	GTC ACG TTT TGG GAA GGG CGA TCG GCG AAA GG
7[32]_9[31]	ATA CCA AGA CAG ACC AGG CGC ATA GAC AAG AA
12[63]_14[64]	GAT AAA ATC ATA AAT ATT CAT TAG CGG ATT
30[47]_28[48]	TGT TTG ATG CGG GGA GAG GCG GTT TTA ATT GC
21[192]_23[191]	ACA AAA TCC CTA CCA TAT CAA AAT TGA TTA TC
0[111]_2[96]	GCG GAT AAG TGC CGT CGA GAG GGT GAG TTT CGT CAC CAG GAC AGC CC
10[95]_12[96]	AAT TTC AAT TCA ACT AAT GCA GAT GTA AGA GC
5[80]_3[79]	GCA TCG GAG TGA ATT TCT TAA ACA TTT TCA CG
2[175]_0[176]	TGG CCT TGG ATA CAG GAG TGT ACT CAT GAA AG
9[32]_11[31]	CCG GAT ATT ATA CCA GTC AGG ACG AAC TAA CG
16[143]_14[144]	GGG CTT AAG TAA TTT AGG CAG AGG TAA TTT AC
5[160]_7[159]	GCC AGC AAT GAG GGA GGG AAG GTA ACG TAG
4[95]_6[96]	CTT TCG AGA CGA GGG TAG CAA CGG TCA TGA GG
22[175]_20[176]	CGT AAA ACC CTG ATT GCT TTG AAT TAC ATT TA
2[207]_0[208]	ATT GAC AGG GGT CAG TGC CTT GCC TGC CTA
29[32]_31[31]	GCC AAC GCG GTG GTT CCG AAA TCG CCG AGA TA
10[47]_8[48]	TGG CTC ATT CAT TAC CCA AAT CAA GAC AGA TG
27[192]_29[191]	CCT AAA ACA CAT TGG CAG ATT CAC CCG CCA GC
23[192]_25[191]	AGA TGA TGA GAA GTA TTA GAC TTT AAG GTT AT
22[63]_24[64]	TAT TTT GGC TTT CAT CAA CAT TTC GTA ACC
2[143]_0[144]	CAT TAA AGG TTC CAG TAA GCG TCA AGA GAA GG
11[160]_13[159]	ACC AAC GCA ACC TCC CGA CTT GCG TAT TAA
1[80]_0[64]	CGT AAC ACT TGA TAT AAG TAT AGC CCG GAA TA
1[32]_3[31]	ACC ACC CTA ATT TTC TGT ATG GGA GGA GTG AG
12[47]_10[48]	AGC GAG AGT TAT TAC AGG TAG AAA TTA AGA AC
16[207]_14[208]	CCT GTT TAA GTA ATT CTG TCC AGC AGA ACG
8[143]_6[144]	GGA ATA CCA CGC AGT ATG TTA GCA AAA TAT TG
1[192]_3[191]	GTT TTA ACG GAG GTT GAG GCA GGT TCA GAA CC
8[95]_10[96]	TCA ATC ATC TTG CCC TGA CGA GAA GAT GGT TT
27[80]_25[79]	TTC CTG TGC TGC AAG GCG ATT AAG TCA GGC TG
12[95]_14[96]	AAC ACT ATC TTT AAA CAG TTC AGA GTC TTT AC
2[63]_4[64]	GTC TTT CTT GCG AAT AAT AAT TGC TTG ATA
28[175]_26[176]	CGA CCA GTG TCT TTA ATG CGC GAA AAC AGT GC
23[32]_25[31]	TCG CGT CTC GAC GAC AGT ATC GGC CGG CAC CG
16[175]_14[176]	CTT ACC AGA ATA AGA GAA TAT AAA AGT CCT GA
22[207]_20[208]	GTT AGA AGC GCA GAG GCG AAT TAT GAA ACA
20[47]_18[48]	AAA TTA ATT ATT TCA ACG CAA GGA TAA AGC CT
28[95]_30[96]	TGT AAA GCT CAC CAG TGA GAC GGG CTG AGA GA

22[95]_24[96]	GTA TAA GCC AAC CCG TCG GAT TCT TGG GAT AG
6[143]_4[144]	ACG GAA ATA CTT GAG CCA TTT GGG TCA TAG CC
31[82]_29[79]	TGG ACT AAG CGG TCC ACG CTG GAG GGT GGT
21[32]_23[31]	TGA GAG TCA TTT TTG TTA AAT CAG AAA ATA AT
17[32]_19[31]	ATA CAT TTA AAG CTA AAT CGG TTG TTG CGG GA
16[95]_18[96]	AGC TCA ACA GCT GAA AAG GTG GCA TCC AAT AA
26[143]_24[144]	TAA AGC ATA TCA ATA TCT GGT CAG TTG CCC GA
15[160]_17[159]	GAG CCA GTT ATA AAG CCA ACG CTA TTT CAT
19[80]_17[79]	TAT TTT AAG GCA AGG CAA AGA ATT TTT CAT TT
15[80]_13[79]	CCT TTT GAT TAT AGT CAG AAG CAA GAA TCC CC
3[160]_5[159]	CCG CCT CCC ATC GGC ATT TTC GGA ATT AGA
31[112]_31[143]	GTC TAT CAG GCG CGT ACT ATG GTT GCT TTG AC
6[47]_4[48]	CAA CCT AAC CGC TTT TGC GGG ATC TTG CGC CG
30[175]_28[176]	CTG TCC ATA TCC AGA ACA ATA TTA CAG TCA CA
28[207]_26[208]	ATT ATT TAT CGC CAT TAA AAA TAG GTG AGG
20[143]_18[144]	AAA CAG TAC TTC TGT AAA TCG TCG GCT GAT GC
9[80]_7[79]	GAA TAA GGA AGG GAA CCG AAC TGA ATC GCC TG
6[63]_8[64]	CCA CTA CAC GGA GAT TTG TAT CCC AAC TTT
20[175]_18[176]	ACA ATT TCC CCT TAG AAT CCT TGA CTT AGG TT
11[32]_13[31]	GAA CAA CAG CTT TTG CAA AAG AAG TAG ACT GG
25[80]_23[79]	CGC AAC TGG GTG TAG ATG GGC GCA AAA TGT GA
5[192]_7[191]	TAG CAA GGC GCC AAA GAC AAA AGG CAT ATA AA
10[143]_8[144]	CGC ATT AGT CAG AGG GTA ATT GAG ATA ATA AC
30[95]_31[111]	GTT GCA GCC CAA CGT CAA AGG GCG AAA AAC C
8[207]_6[208]	GAA AAG TAA AGA CAC CAC GGA ACA TAT GGT
31[32]_30[48]	GGG TTG AGT GTT GTT CCA GTT TGG GAA AAT CC
10[175]_8[176]	TTA CAG AGG ATA ACC CAC AAG AAT AAA GTT AC

Hairpin staples:

23[160]_25[159]	GGA GCG GAC TCG TAT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAT CCT TTG GCA A
3[32]_5[31]	AAT AGA AAA ACA ACC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT CGC CCA CTG CAG GGA
29[192]_31[191]	CAT TGC AAA GTG AGG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC ACC GAG TGG AGC TAA
4[207]_2[208]	AAT CAA GCA GAG CCT CCT CTT TTG AGG AAC AAG TTT TCT TGT ACC ACC CTC CGC CAG C
10[207]_8[208]	TGT TTA AAT AAG AGT CCT CTT TTG AGG AAC AAG TTT TCT TGT CAA GAA ACC TTT TTA A
6[207]_4[208]	TTA CCA GCC GGA AAT CCT CTT TTG AGG AAC AAG TTT TCT TGT CGT CAC CAA GCG ACA G
19[160]_21[159]	TTA ATT TTA TTT GAA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTT ACC TTT ACA ATA A

0[175]_1[159]	TAT TAA GAG GCT GAG ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC TCC TCA TAC ATG G
31[160]_30[176]	GCT TTC CTC GTT AGA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT CAG AGC GAA AAG AGT
14[143]_12[144]	GAG CAT GTA TCA TTC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AGA ACG GGG GAG GTT
14[175]_12[176]	ACA AGA AAC CGC ACT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TCG AGA AGC GAG GCG
16[63]_18[64]	ACG GTG TCC TGT TTT CCT CTT TTG AGG AAC AAG TTT TCT TGT AGC TAT ATA GCA AAA T
15[192]_17[191]	AAA AGG TAA GTA TCA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TGC GTT AGA AAT ACC
7[160]_9[159]	AAA ATA CAA ACC GAG GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAC GCA CGC TAA T
0[41]_1[31]	GGT TTA GTA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GCC ACC CTC AGA GCC
26[175]_24[176]	CAC GCT GAT TGA AAG GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATT GAG GAC AAA CAA
13[192]_15[191]	CCG TTT TTT TAT CAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATA GAT AGT ACC GAC
6[95]_8[96]	AAG TTT CCT GTC GAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT CCG CGA CGC AGA CGG
15[32]_17[31]	GGA AGC AAA TAT AAC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG TTG ATT CAC CAT TAG
18[207]_16[208]	GAG ACT AGA TAA ATT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAG GCG TTA GAA AAA G
13[160]_15[159]	ACC AAG TAA ATA ATA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC CCA TCC CAT TTT C
25[32]_27[31]	CTT CTG GTC GTT GTA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA ACG ACG GAC TCT AGA
6[175]_4[176]	CAA CCG ATA ATC ACC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG TAG CAC CGA CTG TAG
10[63]_12[64]	TGC GAT TGA TTC ATT CCT CTT TTG AGG AAC AAG TTT TCT TGT CAG TTG AGC AGA CGA C
14[207]_12[208]	CGC CTG TAT TTT CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TCG TAG GAA TAT AGA A
18[47]_16[48]	CAG AGC ATC GCA AAT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTG TCA ATA ACT GGA AGT
29[160]_31[159]	CTG GTA ATC ACG CAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT TAA CCG ATA ACG T
9[192]_11[191]	GCC CAA TAC GTC AAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA TGA AAA TAG AGC CTA
20[63]_22[64]	ACC GTT CAT CGA TGT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAC GGT AAA ACG TTA A
1[160]_3[159]	CTT TTG ATA TAT TCA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAC AAA ACC GGA A
9[160]_11[159]	ATC AGA GAA GAA TAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TAA AAA GAA TCT T

3[80]_1[79]	TTG AAA ATA GCG TAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTG ATC TAA ACC ATG TAC
4[63]_6[64]	CCG ATA GGT CAC CCT CCT CTT TTG AGG AAC AAG TTT TCT TGT TCA GCA GCA CGT AAT G
0[63]_2[64]	GGT GTA TAA GCC CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT ATA GGA ACG TTT TGT C
23[80]_21[79]	GCG AGT AAA AAT ATT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAT TGT ATC GTA AAA
28[63]_30[64]	ACT CAC ATG CGT ATT CCT CTT TTG AGG AAC AAG TTT TCT TGT TGG GCG CCT TTG CCC C
13[32]_15[31]	ATA GCG TCC CCG AAA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA CTT CAA AAC CAG ACC
26[95]_28[96]	GGG ATG TGT GAA ATT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTT TAT CCG CGC ATA AAG
19[32]_21[31]	GAA GCC TTG CCG GAG ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG GGT AGC TTC ATT GCC
5[32]_7[31]	GTT AAA GGA ACG AAA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GGC AAA ACA GCG ATT
14[95]_16[96]	CCT GAC TAT AAG AGG TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC ATT TTT GAA TGC TGT
2[47]_0[42]	AGT AAA TGC ATT TTC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG GGA TAG CCA CCG TAC TCA GGA
17[80]_15[79]	GGG GCG CGA TGT TTT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATA TGC ATA ATT GCT
30[207]_28[208]	TAT AAT CCA GGA AAT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAC GCT CAT GAA ATG G
18[95]_20[96]	ATC ATA CAA TGC AAT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTC CTG AGT AAA GGC CGG
18[175]_16[176]	GGG TTA TAC TAA ATT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATG GTT TTA CAA ATT
3[192]_5[191]	GCC ACC CTT TTG CCT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTT AGC GTC AAT TAC CAT
20[95]_22[96]	AGA CAG TCT CAA TCA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TGT ACC CGG AAG ATT
28[143]_26[144]	CAG AGA TAT GGC ACA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA CAA TAT TGA AAA ATC
4[143]_2[144]	CCC TTA TTC GGA ACC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG AGC CAC CTA AAT CCT
22[143]_20[144]	TTT CAG GTT TTT ACA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GGG AGA ATT TAA TGG
24[207]_22[208]	AGG ATT TGC AAT TCT CCT CTT TTG AGG AAC AAG TTT TCT TGT ATC AAT ATA TGG AAG G
18[143]_16[144]	AAA TCC AAA AAT ATA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTT TTA GTT ACA ACA GTA
8[47]_6[48]	AAC GGT GTC GCG AAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAG TAC AGA AGG CAC
2[95]_4[96]	TCA TAG TTC TCC AAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAA GGC TTC AGC TTG

26[47]_24[48]	AGT CAC GAG CCG GAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC CAG GCA ATG CCA GTT
13[80]_11[79]	CTC AAA TGC ATA ACC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTT CGT TTA CAT TTA GGA
0[207]_1[191]	TTT CGG AAC CTA TTA TCC TCT TTT GAG GAA CAA GTT TTC TTG TTT CTG AAA GGT AAT AA
14[63]_16[64]	GCA TCA AAT TAG AGT CCT CTT TTG AGG AAC AAG TTT TCT TGT AGT ACC TTA CTA AAG T
24[143]_22[144]	ACG TTA TTG CGG AAC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA AGA AAC CGC GTA GAT
4[175]_2[176]	CGC GTT TTC TCA GAG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GCC ACC CCA GAC GAT
17[160]_19[159]	CTT CTG ACT AAC TAT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT GTA AAT CTA TTA A
4[47]_2[48]	ACA ATG ACG GAA CAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTT AAA GGA ACA GAC GTT
18[63]_20[64]	TAA GCA ATA AAA ATT CCT CTT TTG AGG AAC AAG TTT TCT TGT TTT TAG AAG ATA TTC A
12[143]_10[144]	TTG AAG CCG CTA CAA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTT TTA TCC TCA GGG AAG
17[192]_19[191]	GAC CGT GTC CTT TTT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA CCT CCG GAA ACA TAG
26[63]_28[64]	CGC CAG GTT CGT AAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TCA TGG TCG TGA GCT A
30[63]_31[81]	AGC AGG CAA CAA GAT CCT CTT TTG AGG AAC AAG TTT TCT TGT GTC CAC TAT TAA AGA ACG
11[80]_9[79]	ATA CCA CAC TTT AAT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TTG TGA ACA TTC AGT
26[207]_24[208]	CGG TCA GTC TTT AGT CCT CTT TTG AGG AAC AAG TTT TCT TGT GAG CAC TAT ACA TTT G
12[175]_10[176]	TTT TAG CGT AAC GAG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTG TCT TTC CAG CAG CCT
8[175]_6[176]	CAG AAG GAT ACA TAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG GTG GCA AGC GAC ATT
29[80]_27[79]	TTT TCT TTC TGG GGT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTC CTA ATG AAT AGC TGT
7[80]_5[79]	ATA AAT TGA TTA AAC GTC CTC TTT TGA GGA ACA AGT TTT CTT GTG GTA AAA TGA AAG ACA
25[160]_27[159]	ATC AAC AGG AGC CAG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GCA AAT TTT GAA T
28[47]_26[48]	GTT GCG CTG GGT ACC GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GCT CGA AGT TTT CCC
21[160]_23[159]	CGG ATT CGA GAA ATA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA GAA ATT ACC AGA A
22[47]_20[48]	CGC ATT AAT GGA GCA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA CAA GAG ATA GCT GAT
19[192]_21[191]	CGA TAG CTA GAA AAC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAT TAA TAC CAA GTT

0[143]_0[112]	ATT AGG ATT AGC GGG GTC CTC TTT TGA GGA ACA AGT TTT CTT GTT TTT GCT CAG TAC CAG
21[80]_19[79]	CTA GCA TGA AAT CAC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TCA ATA TCC CTC ATA
16[47]_14[48]	TTC ATT CCA CTC CAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GGT CAG GAA AGA TTA
31[144]_28[144]	GAG CAC GTT TGT AGC AAT ACT TCT AAG AAC TCT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAA CTA TCC TGG CCA A
24[175]_22[176]	TTC GAC AAA TTA TCA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC ATA TTC CTA TTT GCA
7[192]_9[191]	AGA AAC GCA AGC AGA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GCC GAA CTG AGT TAA
8[63]_10[64]	GAA AGA GCG TAA CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAG CTG CTT TAC CTT A
12[207]_10[208]	GGC TTA TGT TAC AAT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAT AAA CAC GAT TTT T
27[32]_29[31]	GGA TCC CCC ACT GCC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTG CTT TCC AAT GAA TCG
27[160]_29[159]	GGC TAT TAA ATA AAA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTG GAC ATT GGC CTT G
31[192]_30[208]	ACA GGA GGC CGA TTA TCC TCT TTT GAG GAA CAA GTT TTC TTG TAA GGG ATT AGT GTT TT
11[192]_13[191]	ATT TGC CAC CGG TAT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC TAA GAA CCA AGC AAG
20[207]_18[208]	AAC ATC ATA GAT TAT CCT CTT TTG AGG AAC AAG TTT TCT TGT AGA CGC TGA GGT CTG A
24[47]_22[48]	TGA GGG GAG GCC TTC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTT GTA GCC ATT AAA ATT
24[63]_26[64]	GTG CAT CAG CGC CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TTC GCC ATT TGG GTA A
14[47]_12[48]	AGA GGA AGC AAT ACT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GGA ATC GAC CAA AAT
25[192]_27[191]	CTA AAA TAT ATT AAC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC CGC CTG CCT GAT AGC
25[128]_26[112]	AAC CCT CAC ACC TTG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTT GAA CCT CCT ATT ACG
6[111]_5[127]	GAC TTT TCT ACA GAT CCT CTT TTG AGG AAC AAG TTT TCT TGT GGC TTT GAA TTA TCA C
15[128]_16[112]	GCC AAC ATT TGA GAA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GCC ATA TTT AAT TGC
24[111]_23[127]	ACC GTA ACC GTG GGT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAC AAA CGG TAA CAT T
13[128]_14[112]	CTT TCC TTA GAA ACC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA TCA ATA ATA AAT CAA
8[111]_7[127]	ACG AGG CCT GCT CCT CCT CTT TTG AGG AAC AAG TTT TCT TGT ATG TTA CTA TTA AGA C
9[128]_10[112]	GAA CAA AGA CGG GAG ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA TTA ACT GGT AAA TTG

14[111]_13[127]	AAA TCA GAA ACG AGT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAT GAC CAT CGG CTG T
12[111]_11[127]	AGG CAT AAC ATA ACT CCT CTT TTG AGG AAC AAG TTT TCT TGT GCC AAA AGT GCT ATT T
22[111]_21[127]	AAA AAC ACG GTT GAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TAA TCA GAA TAC AGT A
16[111]_15[127]	TGA ATA TCG GAT GGT CCT CTT TTG AGG AAC AAG TTT TCT TGT CTT AGA GCT TAA CAA C
3[128]_4[112]	AAA ATC ACA GCG TTT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTC CAT CTT TAA TTG TAT
23[128]_24[112]	ATC ATT TTA ATT TTA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA AGT TTG AGC GGA TTG
5[128]_6[112]	CGT CAC CGT ATT CAT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAG GTG AGG ACT AAA
2[111]_1[127]	TTC CAC ATA CAA ACT CCT CTT TTG AGG AAC AAG TTT TCT TGT TAC AAC GCA GTC TCT G
1[128]_2[112]	AAT TTA CCC CAG AAT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTG AAA GCG CCT GTA GCA
26[111]_25[127]	CCA GCT GGT GCG GGT CCT CTT TTG AGG AAC AAG TTT TCT TGT CCT CTT CGA AAT ATC A
21[128]_22[112]	ACA GTA CCT TAA CGT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GAT GAA TAA AGC CCC
10[111]_9[127]	GGC TTG AAC ACC AGT CCT CTT TTG AGG AAC AAG TTT TCT TGT AAC GAG TAA ACA CCC T
11[128]_12[112]	TGC ACC CAT TAA ATC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA GAT TAG TGA ATT ACG
4[111]_3[127]	CGG TTT ACC AAA AGT CCT CTT TTG AGG AAC AAG TTT TCT TGT GAG CCT TTT CAT AAT C
29[128]_30[112]	CTG AGT AGT TGA TTA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTT AAT AAC ACT TCA CCG
30[111]_29[127]	CCT GGC CCA ACA GCT CCT CTT TTG AGG AAC AAG TTT TCT TGT TGA TTG CCT CAC TTG C
17[128]_18[112]	ACT TTT TCT CGC AAG ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC AAA GAA CGT AGT AGC
18[111]_17[127]	ATT AAC ATC AAT TCT CCT CTT TTG AGG AAC AAG TTT TCT TGT TAC TAA TAG CGA GAA A
27[128]_28[112]	AGA ATA CGG AAC CCT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC TGA CCT GAC ATA CGA
7[128]_8[112]	TCC TTA TTC AAA AGA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC TGG CAT GTA GCC GGA
20[111]_19[127]	GGT GAG AAT GTG TAT CCT CTT TTG AGG AAC AAG TTT TCT TGT GGT AAA GAT GAG TGA A
28[111]_27[127]	GCC GGA ATC ACA ATT CCT CTT TTG AGG AAC AAG TTT TCT TGT TCC ACA CAA AAG CGT A
19[128]_20[112]	TAA CCT TGC ATA AAT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATA TAT GTT CAA AAG
24[95]_26[96]	GTC ACG TTT TGG GAA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTG GCG ATC GGC GAA AGG

7[32]_9[31]	ATA CCA AGA CAG ACC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG GCG CAT AGA CAA GAA
12[63]_14[64]	GAT AAA ATC ATA AAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TAT TCA TTA GCG GAT T
30[47]_28[48]	TGT TTG ATG CGG GGA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GGC GGT TTT AAT TGC
21[192]_23[191]	ACA AAA TCC CTA CCA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA TCA AAA TTG ATT ATC
0[111]_2[96]	GCG GAT AAG TGC CGT CGA GAG GGT GAG TTT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTG TCA CCA GGA CAG CCC
10[95]_12[96]	AAT TTC AAT TCA ACT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA TGC AGA TGT AAG AGC
5[80]_3[79]	GCA TCG GAG TGA ATT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC TTA AAC ATT TTC ACG
2[175]_0[176]	TGG CCT TGG ATA CAG GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GTG TAC TCA TGA AAG
9[32]_11[31]	CCG GAT ATT ATA CCA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTT CAG GAC GAA CTA ACG
16[143]_14[144]	GGG CTT AAG TAA TTT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG GCA GAG GTA ATT TAC
5[160]_7[159]	GCC AGC AAT GAG GGA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTG GAA GGT AAC GTA G
4[95]_6[96]	CTT TCG AGA CGA GGG TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GCA ACG GTC ATG AGG
22[175]_20[176]	CGT AAA ACC CTG ATT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTC TTT GAA TTA CAT TTA
2[207]_0[208]	ATT GAC AGG GGT CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT GTG CCT TGC CTG CCT A
29[32]_31[31]	GCC AAC GCG GTG GTT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GAA ATC GCC GAG ATA
10[47]_8[48]	TGG CTC ATT CAT TAC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC AAA TCA AGA CAG ATG
27[192]_29[191]	CCT AAA ACA CAT TGG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GAT TCA CCC GCC AGC
23[192]_25[191]	AGA TGA TGA GAA GTA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTT AGA CTT TAA GGT TAT
22[63]_24[64]	TAT TTT GGC TTT CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TCA ACA TTT CGT AAC C
2[143]_0[144]	CAT TAA AGG TTC CAG TTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AGC GTC AAG AGA AGG
11[160]_13[159]	ACC AAC GCA ACC TCC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTG ACT TGC GTA TTA A
1[80]_0[64]	CGT AAC ACT TGA TAT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA GTA TAG CCC GGA ATA
1[32]_3[31]	ACC ACC CTA ATT TTC TTC CTC TTT TGA GGA ACA AGT TTT CTT GTG TAT GGG AGG AGT GAG
12[47]_10[48]	AGC GAG AGT TAT TAC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG GTA GAA ATT AAG AAC

16[207]_14[208]	CCT GTT TAA GTA ATT CCT CTT TTG AGG AAC AAG TTT TCT TGT TCT GTC CAG CAG AAC G
8[143]_6[144]	GGA ATA CCA CGC AGT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT GTT AGC AAA ATA TTG
1[192]_3[191]	GTT TTA ACG GAG GTT GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GGC AGG TTC AGA ACC
8[95]_10[96]	TCA ATC ATC TTG CCC TTC CTC TTT TGA GGA ACA AGT TTT CTT GTG ACG AGA AGA TGG TTT
27[80]_25[79]	TTC CTG TGC TGC AAG GTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GAT TAA GTC AGG CTG
12[95]_14[96]	AAC ACT ATC TTT AAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GTT CAG AGT CTT TAC
2[63]_4[64]	GTC TTT CTT GCG AAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TAA TAA TTG CTT GAT A
28[175]_26[176]	CGA CCA GTG TCT TTA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT GCG CGA AAA CAG TGC
23[32]_25[31]	TCG CGT CTC GAC GAC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTG TAT CGG CCG GCA CCG
16[175]_14[176]	CTT ACC AGA ATA AGA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATA TAA AAG TCC TGA
22[207]_20[208]	GTT AGA AGC GCA GAT CCT CTT TTG AGG AAC AAG TTT TCT TGT GGC GAA TTA TGA AAC A
20[47]_18[48]	AAA TTA ATT ATT TCA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC GCA AGG ATA AAG CCT
28[95]_30[96]	TGT AAA GCT CAC CAG TTC CTC TTT TGA GGA ACA AGT TTT CTT GTG AGA CGG GCT GAG AGA
22[95]_24[96]	GTA TAA GCC AAC CCG TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GGA TTC TTG GGA TAG
6[143]_4[144]	ACG GAA ATA CTT GAG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC ATT TGG GTC ATA GCC
31[82]_29[79]	TGG ACT AAG CGG TCT CCT CTT TTG AGG AAC AAG TTT TCT TGT CAC GCT GGA GGG TGG T
21[32]_23[31]	TGA GAG TCA TTT TTG TTC CTC TTT TGA GGA ACA AGT TTT CTT GTT AAA TCA GAA AAT AAT
17[32]_19[31]	ATA CAT TTA AAG CTA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA TCG GTT GTT GCG GGA
16[95]_18[96]	AGC TCA ACA GCT GAA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA GGT GGC ATC CAA TAA
26[143]_24[144]	TAA AGC ATA TCA ATA TTC CTC TTT TGA GGA ACA AGT TTT CTT GTC TGG TCA GTT GCC CGA
15[160]_17[159]	GAG CCA GTT ATA AAG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC AAC GCT ATT TCA T
19[80]_17[79]	TAT TTT AAG GCA AGG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAG AAT TTT TCA TTT
15[80]_13[79]	CCT TTT GAT TAT AGT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA GAA GCA AGA ATC CCC
3[160]_5[159]	CCG CCT CCC ATC GGC ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT TTT CGG AAT TAG A

31[112]_31[143]	GTC TAT CAG GCG CGT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC TAT GGT TGC TTT GAC
6[47]_4[48]	CAA CCT AAC CGC TTT TTC CTC TTT TGA GGA ACA AGT TTT CTT GTG CGG GAT CTT GCG CCG
30[175]_28[176]	CTG TCC ATA TCC AGA ATC CTC TTT TGA GGA ACA AGT TTT CTT GTC AAT ATT ACA GTC ACA
28[207]_26[208]	ATT ATT TAT CGC CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TTA AAA ATA GGT GAG G
20[143]_18[144]	AAA CAG TAC TTC TGT ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA ATC GTC GGC TGA TGC
9[80]_7[79]	GAA TAA GGA AGG GAA CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC GAA CTG AAT CGC CTG
6[63]_8[64]	CCA CTA CAC GGA GAT CCT CTT TTG AGG AAC AAG TTT TCT TGT TTT GTA TCC CAA CTT T
20[175]_18[176]	ACA ATT TCC CCT TAG ATC CTC TTT TGA GGA ACA AGT TTT CTT GTA TCC TTG ACT TAG GTT
11[32]_13[31]	GAA CAA CAG CTT TTG CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAA GAA GTA GAC TGG
25[80]_23[79]	CGC AAC TGG GTG TAG ATC CTC TTT TGA GGA ACA AGT TTT CTT GTT GGG CGC AAA ATG TGA
5[192]_7[191]	TAG CAA GGC GCC AAA GTC CTC TTT TGA GGA ACA AGT TTT CTT GTA CAA AAG GCA TAT AAA
10[143]_8[144]	CGC ATT AGT CAG AGG GTC CTC TTT TGA GGA ACA AGT TTT CTT GTT AAT TGA GAT AAT AAC
30[95]_31[111]	GTT GCA GCC CAA CGT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA AAG GGC GAA AAA CC
8[207]_6[208]	GAA AAG TAA AGA CAT CCT CTT TTG AGG AAC AAG TTT TCT TGT CCA CGG AAC ATA TGG T
31[32]_30[48]	GGG TTG AGT GTT GTT CTC CTC TTT TGA GGA ACA AGT TTT CTT GTC AGT TTG GGA AAA TCC
10[175]_8[176]	TTA CAG AGG ATA ACC CTC CTC TTT TGA GGA ACA AGT TTT CTT GTA CAA GAA TAA AGT TAC

4: Linker edge staple strands

Tile 1:

tile_1_1_1_5l	ATGGACACCTGCCTAAACAGTTAATGCCCAGTAACAGTGCCCGTATTTTGGCAGGTG
tile_1_1_5_3l	CTACTCCGTTTTCAGCACCGTAATCAGTATGAAACCATCGATAGCGGAGTAGTCCAT
tile_1_1_9_3l	GTCTACGGTTTTCTATCTTACCGAAGCCAATGAAATAGCAATAGCCGTAGACGGTAT
tile_1_23_5l_beta	CCTGTGGAGATCCATAGCAAGCAAATCAGATCATTACCGCGCCCATTTTGGATCTCC

tile_1_26_5l_beta	AAGTCGATGCGGTAATTTATCAAAATCATAGAAGAGTCAATAGTGTTTTACCGCATC
tile_1_1_27_5l	AGAGTCACGAGGTGCAGAAGATAAAACAGACCGAACGAACCACCATTTTACCTCGTG
tile_1_28_3l_beta	CACGTCCATTTTATTATACTTCTGAATAAATCCTGATTGTTTGGTGGACGTGAGAGT
tile_1_1_31_3l	GTCTGCCGTTTTCGCCAGAATCCTGAGATTAGACAGGAACGGTACGGCAGACGTTGA

Tile 2:

tile_2_23_5l_beta	TCCATCCGCTCCGATAGCAAGCAAATCAGATCATTACCGCGCCCATTTTCGGAGCGG
tile_2_1_23_5l	CTAACGAGATGCCATCAAGAGTAATCTTGGCTGGCTGACCTTCTTTTGGCATCTC
tile_2_1_27_5l	CTGAGACCAGTCCTCGGTCGCTGAGGCTGCATAACCGATATATTTTTGGACTGGT
tile_2_1_31_3l	GGTAGCCATTTTTCAGAACCGCCACCCTCAGAACCGCCACCCTGGCTACCTTGTC
tile_2_7_3l_beta	CCGTGTCGTTTTTAATAGTAAAATGTTTTTTGCCAGAGGGGGCGACACGGCAGTA
tile_2_0_24_3l	CTACACCGTTTTCTATCTTACCGAAGCCAATGAAATAGCAATAGCGGTGTAGATGGA
tile_2_0_28_3l	GGTGGAAGTTTTCAGCACCGTAATCAGTATGAAACCATCGATAGCTTCCACCATACC
tile_2_0_32_5l	ACAGGGCTCGTCCTAAACAGTTAATGCCCAGTAACAGTGCCCGTATTTTGGACGAGC

Tile 3:

tile_3_1_5l_beta	GACTTGTCCGTCCTCAGAACCGCCACCCTCAGAACCGCCACCCTTTTGGACGGAC
tile_3_3_3l_beta	CACAGCAGTTTTTCGGTCGCTGAGGCTGCATAACCGATATATCTGCTGTGACTCT
tile_3_5_5l_beta	ACTCTCGTAGTCGATCAAGAGTAATCTTGGCTGGCTGACCTTCTTTTCGACTACG
tile_3_7_3l_beta	GCGGATCGTTTTTAATAGTAAAATGTTTTTTGCCAGAGGGGGCGATCCGCTCAAC
tile_3_17_5l_beta	TTACCCTACCACGTAAACAGTTAATGCCCAGTAACAGTGCCCGTATTTTCGTGGTAG
tile_3_19_3l_beta	GTAGGTCATTTTCAGCACCGTAATCAGTATGAAACCATCGATAGTGACCTACCCTGT
tile_3_21_3l_beta	TCCGTGGCTTTTCTATCTTACCGAAGCCAATGAAATAGCAATAGGCCACGGACTGAG

tile_3_23_5l_beta	AACTGGATACGACATAGCAAGCAAATCAGATCATTACCGCGCCCATTTTGTCGTATC

Tile 4:

tile_4_0_2_3l	CCTTGGAGTTTTCGCCAGAATCCTGAGATTAGACAGGAACGGTACTCCAAGGTACTG
tile_4_0_6_5l	GTTAGGTGCTCTCGCAGAAGATAAAACAGACCGAACGAACCACCATTTTGAGAGCAC
tile_4_0_10_5l	CTCAGACAGCACGATTATACTTCTGAATAAATCCTGATTGTTTGGTTTTCGTGCTGT
tile_4_26_3l_beta	GAGCGACGTTTTAATTTATCAAAATCATAGAAGAGTCAATAGTGCGTCGCTCGACAA
tile_4_23_5l_beta	GGTAACCAGTTCGATAGCAAGCAAATCAGATCATTACCGCGCCCATTTTCGAACTGG
tile_4_21_3l_beta	GAGCGTAGTTTTCTATCTTACCGAAGCCAATGAAATAGCAATAGCTACGCTCACAGG
tile_4_0_28_3l	TGACCTCGTTTTCAGCACCGTAATCAGTATGAAACCATCGATAGCGAGGTCACTCAG
tile_4_0_32_5l	CAGTTGTGCCTTGTAAACAGTTAATGCCCAGTAACAGTGCCCGTATTTTCAAGGCAC

5: Blocker edge staple strands

List of All Blockers:

Tile_1_1_bl	TCAGAACCGCCACCCTCAGAACCGCCACCCTTTT
Tile_1_2_bl	TTCAACAGTTTCAGCTTTTGCTAAACAACTTTTT
Tile_1_3_bl	TCGGTCGCTGAGGCTGCATAACCGATATATTTTT
Tile_1_4_bl	CTCATCTTTGACCCCGAATACACTAAAACATTTT
Tile_1_5_bl	ATCAAGAGTAATCTTGGCTGGCTGACCTTCTTTT
Tile_1_6_bl	TACGTTAATAAAACGTTGGGAAGAAAAATCTTTT
Tile_1_7_bl	TAATAGTAAAATGTTTTTGCCAGAGGGGGTTTT
Tile_1_8_bl	CGAGCTTCAAAGCGATATCGCGTTTTAATTTTTT
Tile_1_9_bl	AGTAGATTTAGTTTGCCAATTCTGCGAACGTTTT
Tile_1_10_bl	GACCCTGTAATACTTTACCAAAAACATTATTTTT
Tile_1_11_bl	ACAAAGGCTATCAGGATTTTTGAGAGATCTTTTT
Tile_1_12_bl	ATAGGAACGCCATCACTCATTTTTTAACCATTTT
Tile_1_13_bl	CTCCAGCCAGCTTTCCTCAGGAAGATCGCATTTT
Tile_1_14_bl	CATGCCTGCAGGTCGCCAGTGCCAAGCTTGTTTT
Tile_1_15_bl	GTGCCAGCTGCATTAGTCGGGAAACCTGTCTTTT
Tile_1_16_bl	AATCAAAAGAATAGCGCAAAATCCCTTATATTTT

Tile_1_17_bl	TAAACAGTTAATGCCCAGTAACAGTGCCCGTATTTT
Tile_1_18_bl	ACCACCACCAGAGCCGCAGAGCCGCCACCAGATTTT
Tile_1_19_bl	CAGCACCGTAATCAGTATGAAACCATCGATAGTTTT
Tile_1_20_bl	CAATCAATAGAAAATTTAAGTTTATTTTGTCATTTT
Tile_1_21_bl	CTATCTTACCGAAGCCAATGAAATAGCAATAGTTTT
Tile_1_22_bl	CAATCCAAATAAGAAAGCCATATTATTTATCCTTTT
Tile_1_23_bl	ATAGCAAGCAAATCAGATCATTACCGCGCCCATTTT
Tile_1_24_bl	AACATGTTCAGCTAATGACGACGACAATAAACTTTT
Tile_1_25_bl	CGGAATCATAATTACTAAATAAGAATAAACACTTTT
Tile_1_26_bl	AATTTATCAAAATCATAGAAGAGTCAATAGTGTTTT
Tile_1_27_bl	TGAGCAAAAGAAGATGATTCATTTCAATTACCTTTT
Tile_1_28_bl	ATTATACTTCTGAATAAATCCTGATTGTTTGGTTTT
Tile_1_29_bl	AGCCGTCAATAGATAAACAACTAATAGATTAGTTTT
Tile_1_30_bl	GCAGAAGATAAAACAGACCGAACGAACCACCATTTT
Tile_1_31_bl	TTGACGCTCAATCGTCTGGAAATACCTACATTTTTT
Tile_1_32_bl	CGCCAGAATCCTGAGATTAGACAGGAACGGTATTTT

Exchange these blockers for Linker Edge Staples:

Tile 1: [T2 binding] 17, 19, 21, 23; [T3 Binding] 26, 28, 30, 32

Tile 2: [T1 binding] 17, 19, 21, 23; [T4 Binding] 1, 3, 5, 7

Tile 3: [T1 binding] 1, 3, 5, 7; [T4 Binding] 17, 19, 21, 23

Tile 4: [T2 binding] 26, 28, 30, 32; [T3 Binding] 17, 19, 21, 23

2.6 Acknowledgments

The authors would like to thank J. Zenk for component sequence design and assistance with AFM measurements, A. M. Mohammed for assistance in the fluorophore labeling and A. Cangialosi for helpful discussions and advice on the manuscript. This work was supported by NSF grants CCF-1161941 and CMMI-1253876. We also acknowledge DOE BMES grant DE-SC0010595 for some materials and supplies.

CHAPTER 3: DNA STRAND-DISPLACEMENT TIMER CIRCUITS

Summary

Chemical circuits can coordinate elaborate sequences of events in cells and tissues, from the self-assembly of biological complexes to the sequence of embryonic development. However, autonomously directing the timing of events in synthetic systems using chemical signals remains challenging. Here we demonstrate that a simple synthetic DNA strand-displacement circuit can release target sequences of DNA into solution at a constant rate after a tunable delay that can range from hours to days. The rates of DNA release can be tuned to the order of 1−100 nM per day. Multiple timer circuits can release different DNA strands at different rates and times in the same solution. This circuit can thus facilitate precise coordination of chemical events in vitro without external stimulation.

3.1 Introduction

While gene networks in cells can orchestrate intricate processes by modulating gene expression to release a series of target molecules at specified times,[178,179] synthetic *in vitro* biochemical protocols commonly involve manual steps performed by an experimenter, in which reagents are added, filtered, heated or otherwise altered. Artificial mechanisms to automate the temporal release of trigger molecules would make it possible to direct sequential events without the need for external stimulation. Further, the timed release of molecules using such a process could act as a trigger to control acellular self-assembly processes,[91,180,181] multistep reaction cascades,[25] or to time the release of signaling molecules or other reagents in cell culture.[182–184]

In this paper, we build a chemical *timer circuit* that releases a target sequence of DNA at a constant rate from DNA complexes after a tunable delay period. In contrast to previously designed synthetic *in vitro* transcriptional timer circuits,[185] our timer is controlled solely by DNA strand-displacement processes, which have previously been used to perform diverse information processing tasks including Boolean logic,[87,107,108,186,187] signal amplification,[109–111,188] neural network computation[97] and oscillatory signal generation.[189] The timer circuit is designed such that the strand that is released can be coupled to many of these systems in their present form, suggesting that timer circuits will make it possible to activate elaborate information processing tasks at specified times. Furthermore, a design based on strand displacement reactions alone should allow the circuit to operate in a variety of buffers and at a variety of temperatures without redesign.[109]

3.2 Results and Discussion

The timer circuit works by suppressing the release of a single-stranded DNA molecule for a delay period, after which the molecule is allowed to increase in concentration at a constant rate. We show how to design a timer circuit within an abstract chemical reaction network, and then describe an implementation of the abstract network using a simple set of DNA molecules that interact through strand-displacement reactions. Next, we investigate the range of delay periods and release rates that are possible using our circuit, and finally demonstrate that multiple timer circuits can operate within the same solution.

The timer circuit consists of two simultaneous abstract chemical processes: *production* (Eqn. 3.1) and *delay* (Eqn. 3.2)

$$\emptyset_{Source} \xrightarrow{k_{\text{prod}}} \mathbf{O} \tag{3.1}$$

$$\mathbf{O} + \mathbf{D} \xrightarrow{k_{\text{delay}}} \emptyset_{Waste} \tag{3.2}$$

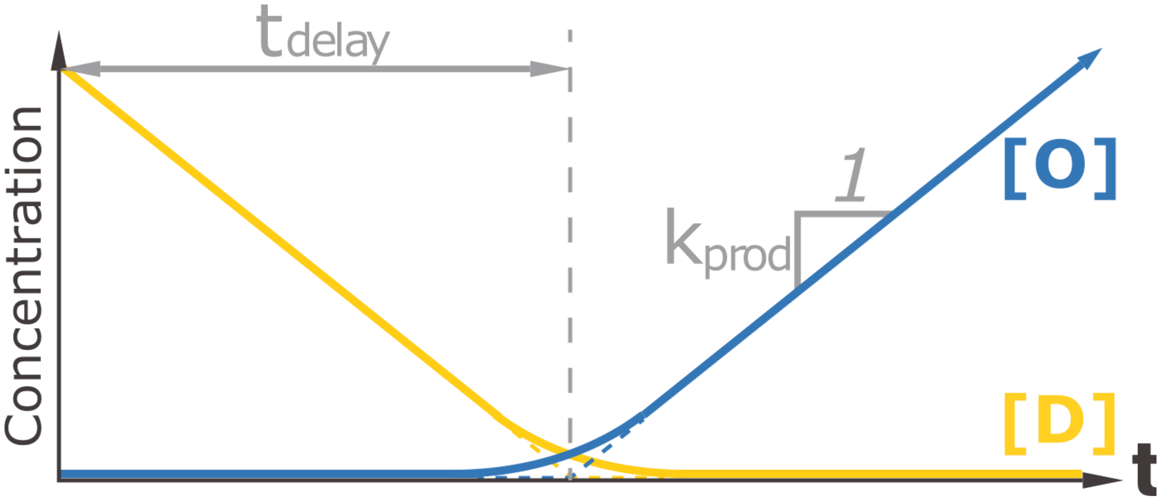


Figure 3.1: Schematic for the operation of a timer circuit. The output species (blue) is constantly produced at rate k_{prod}, but is rapidly consumed by the delay species (yellow). This rapid consumption prevents the accumulation of output until time t_{delay}, when the delay species is depleted.

In the production process, the output **O** is released by a zero-order reaction at a constant rate k_{prod}. In the delay reaction, **O** is rapidly converted into inert waste when it reacts with a delay species **D**. If $k_{\text{delay}}[\mathbf{O}][\mathbf{D}] >> k_{\text{prod}}$, **O** cannot accumulate until all of **D** has been depleted (Fig. 3.1).

We call the time during which **O** cannot accumulate the delay time, t_{delay}, which is the time needed to produce enough **O** to consume all of the **D** that is initially present:

$$\mathrm{t}_{\text{delay}} = \frac{[\mathbf{D}]_0}{k_{\text{prod}}} \tag{3.3}$$

This time can be easily tuned by changing the initial concentration of **D**.

During the delay period, [**O**] remains very small because any molecules of **O** that are produced are rapidly removed. After **D** is depleted, however, [**O**] increases linearly with time. The approximate concentration of free **O** is therefore

$$[\mathbf{O}] \approx \begin{cases} \dfrac{k_{\text{prod}}}{k_{\text{delay}}[\mathbf{D}]} \approx 0, & \text{if } \mathrm{t} < \mathrm{t}_{\text{delay}} \\ k_{\text{prod}}(\mathrm{t} - \mathrm{t}_{\text{delay}}), & \text{otherwise} \end{cases} \tag{3.4}$$

where [**O**] and [**D**] are functions of time.

To construct a timer circuit that controls the release of a DNA strand, we built a set of DNA strand-displacement reactions that emulate the abstract reactions in Equations 3.1 and 3.2 (Fig. 3.2).[190,191] The domain level structure of our complexes follows a DNA architecture previously used for Boolean logic circuit evaluation.[108] Within this implementation, strand **O** is initially partially bound within a complex. The production process frees **O** from this complex, making **O** available in its full single-stranded form (Fig. 3.2a). The delay process likewise sequesters **O** in a waste complex in which the toehold domain of **O** is covered (Fig. 3.2b). Because an exposed toehold domain is generally required to initiate downstream strand-displacement reactions, the delay circuit will control when **O** is available in a functional form.

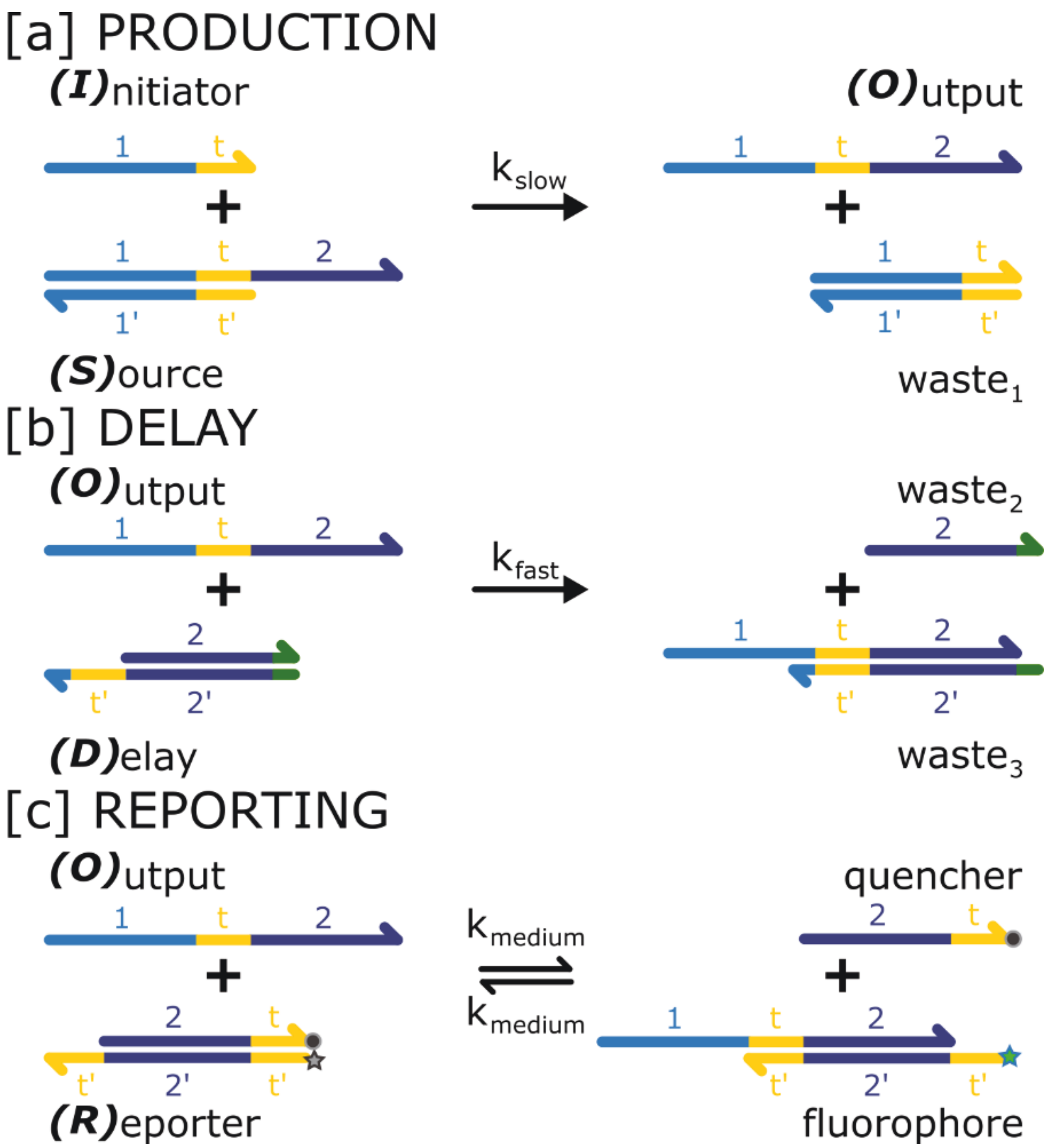


Figure 3.2: Strand-displacement reactions for a timer circuit. *(a) PRODUCTION:* Output is slowly released from source in the presence of initiator. *(b) DELAY:* Output is rapidly consumed by the delay complex. *(c) REPORTING:* Free output binds reversibly to a reporter complex, separating quencher and fluorophore modifiers. FAM and TexasRed paired with appropriate quenchers were used to report on two different output sequences (Section 3.4.0).

Production

The production reaction releases an output molecule **O** when a source complex **S** and an initiator strand **I** react (Fig. 3.2a). This strand-displacement process is facilitated by the spontaneous pairing and unpairing of the bases on the ends of the source complex, *i.e.*, fraying. During the short time periods when these bases are frayed, **I** can bind and compete with **O** until one or the other is displaced.

Reactions initiated by fraying alone have an exceptionally small reaction rate constant, on the order of 0.5 $M^{-1}s^{-1}$, which we denote by k_{0bp}.[96,98] On the order of days, very little **S** and **I** react, allowing us to assume [**S**] and [**I**] remain effectively constant when considering shorter time scales. To release an appreciable concentration of **O**, we use a large amount of **S** and **I** in a reaction. The rate at which **O** is released into solution can therefore be approximated as a constant we term k_{prod}:

$$\frac{\mathrm{d}[\mathbf{O}]}{\mathrm{dt}} = k_{0\mathrm{bp}}[\mathbf{S}][\mathbf{I}] \approx k_{\mathrm{prod}} \qquad (3.5)$$

Equation 3.5 shows that k_{prod} can be easily tuned by changing the initial concentrations of **S** and **I**. For simplicity, and to maximize the time during which the approximation of constant concentrations is reasonable, we keep their initial concentrations equal, *i.e.*, $[\mathbf{S}]_0=[\mathbf{I}]_0$. At longer time scales, the approximation of a constant k_{prod} is violated and [**O**] increases according to second order reaction kinetics.

Delay

To keep [**O**] low while the delay species is present, the delay reaction must sequester **O** at a rate much faster than k_{prod}. The delay complex **D** (Fig. 3.2b), which has a 7 base pair (7 bp) single-stranded toehold domain that binds to **O** and colocalizes it with **D**, acts as a concentration thresholding device.[87] Reactions mediated by 7 bp toeholds proceed at approximately 6 orders of magnitude faster than reactions without a mediating toehold.[96,98] A 2 bp clamp (green in Fig. 3.2b) inhibits some undesired interactions between **S** and **D** while ensuring that the reaction between **O** and **D** remains strongly forward-biased. Clamps with only 1 bp may not reliably prevent interactions at that end and clamps with greater numbers of bases (*e.g.*, 5 bp) are expected to strongly decrease the sequestering ability of the thresholding device due to reaction reversibility.[107,109]

To monitor the reaction's progress, we also include a reporter complex modified with a fluorophore and an associated quencher to track the concentration of free **O** over time. This complex reacts reversibly with the output strand on a time scale much faster than the production reaction, but slower than the delay process, and produces fluorescence as a function of [**O**] at a given time (Fig. 3.2c). The concentration of **O** is related to the fluorescence levels using a calibration curve (Section 3.4.1).[192,193] To build a timer circuit, the source complex **S** and the delay complex **D** are initially combined and the timer is triggered upon the addition of the initiator strand **I**.

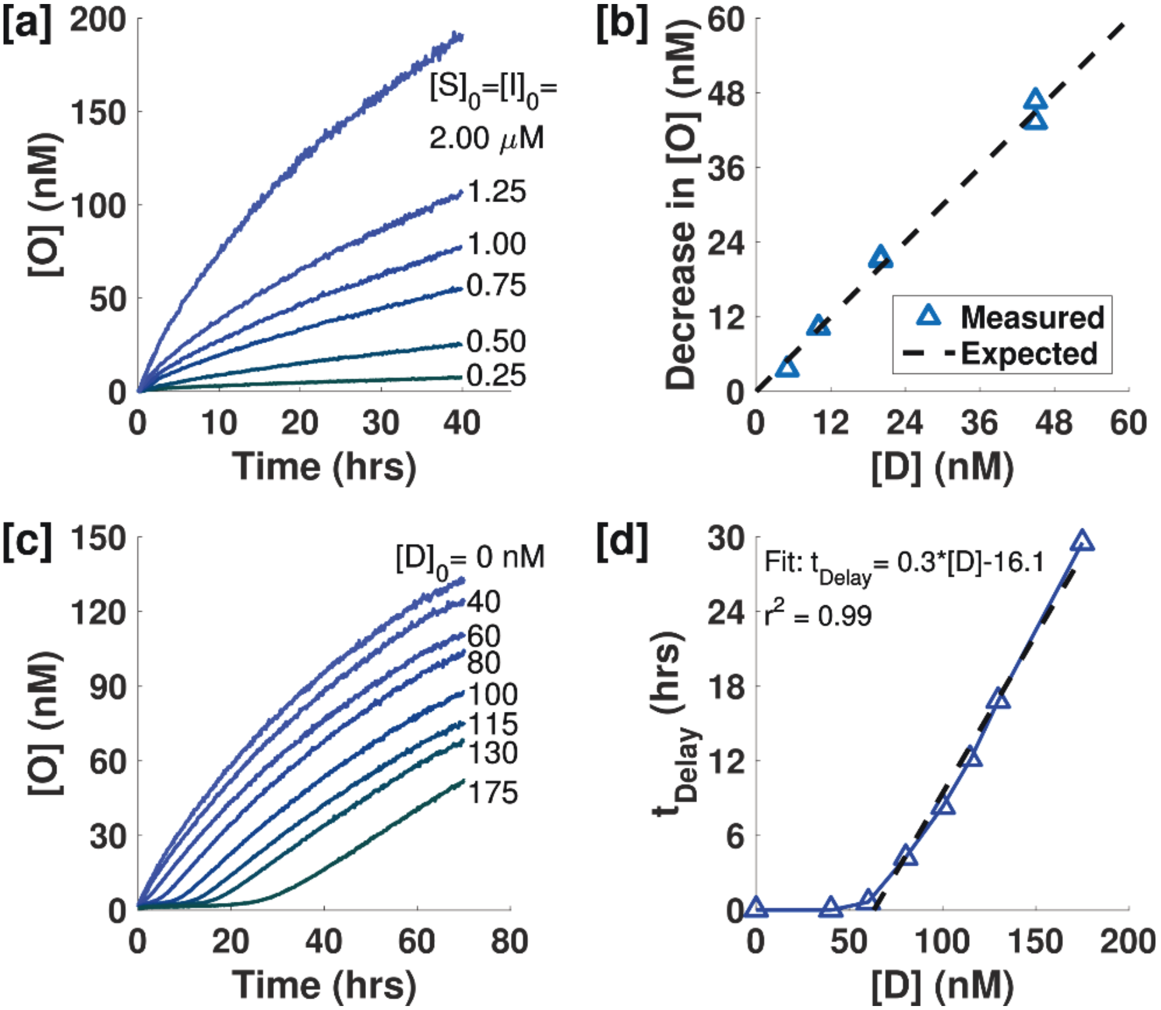


Figure 3.3: Production, Delay and Timer circuit reactions. *(a)* Production kinetics in the absence of **D**. Release rate is dependent upon initial [**S**] and [**I**]. *(b)* [**O**] decreases proportionally to added [**D**]. *(c)* Delay time was tuned by changing the initial concentration of the delay species **D**. [**S**]$_0$=[**I**]$_0$=1 μM. *(d)* Plot of delay time *vs.* initial concentration of **D**, showing an approximately linear relationship after an initial offset.

To demonstrate that the individual reactions performed as desired and determine how the production rate varied with initial concentrations of **S** and **I**, we tested each reaction in isolation at 25 °C (Fig. 3.3a–b, Section 3.4.2). To test the production reaction, we varied [**S**]$_0$=[**I**]$_0$ from 0.25 μM to 2 μM (Fig. 3.3a) and determined the average k_{0bp} to be 0.49±0.13 $M^{-1}s^{-1}$ (Table S3.2) which is in good agreement with previous estimates.[96,98] We calculated, using [**S**]$_0$ and [**I**]$_0$, that k_{prod} varied from 0.15 to 4 nM/hour over the range of concentrations tested

(Table S3.2). When the delay reaction was tested in isolation, the delay complex sequestered free **O**, resulting in a sudden decrease in output detected by the reporter (Fig. S3.7). The decrease in [**O**] matched the concentration of **D** added to the reaction solution (Fig. 3.3b).

To characterize the delay time before **S** begins accumulating as a function of $[\mathbf{D}]_0$ (Eqns. 3.3 and 3.4), we varied $[\mathbf{D}]_0$ while keeping the production rate (determined by $[\mathbf{S}]_0$ and $[\mathbf{I}]_0$) constant. **O** remained low for a delay period that increased with $[\mathbf{D}]_0$ (Fig. 3.3c). For each trial, we used linear least-squares fitting to identify the portion of the production regime with the steepest slope (Section 3.4.3) and used the slope and y-intercept of this fit to calculate the delay time. This method allowed us to measure the delay time without being affected by the ramp up in release that occurs because small amounts of **D** are still present when release noticeably begins. We observed that the delay time varied linearly with respect to the initial concentration of **D** (Fig. 3.3d). We also tested two other timer circuits with lower initial concentrations of **S**, **I** and **D**, and obtained systems with similar delay times but slower rates of output release (Section 3.4.3).

Because the delay circuit is based on DNA strand-displacement events involving a particular DNA sequence, it is possible to create multiple circuits that use different sequences and can trigger the delayed release of two different DNA strands (Fig. 3.4). To characterize the operation of two timer circuits in a single solution, we prepared a second timer circuit and reporter complex with different sequences and fluorophore/quencher than our original system. The second system had the same qualitative behavior, and it was possible to programmatically tune both delay times and release rates. Differences in rates and delay times were observed between the two systems, possibly due to differences in toehold sequence that affect reaction rates (Section 3.4.4–3.4.5). The two systems were able to operate together in the same solution with virtually identical kinetics to those observed when the systems were operated apart (Fig. 3.4).

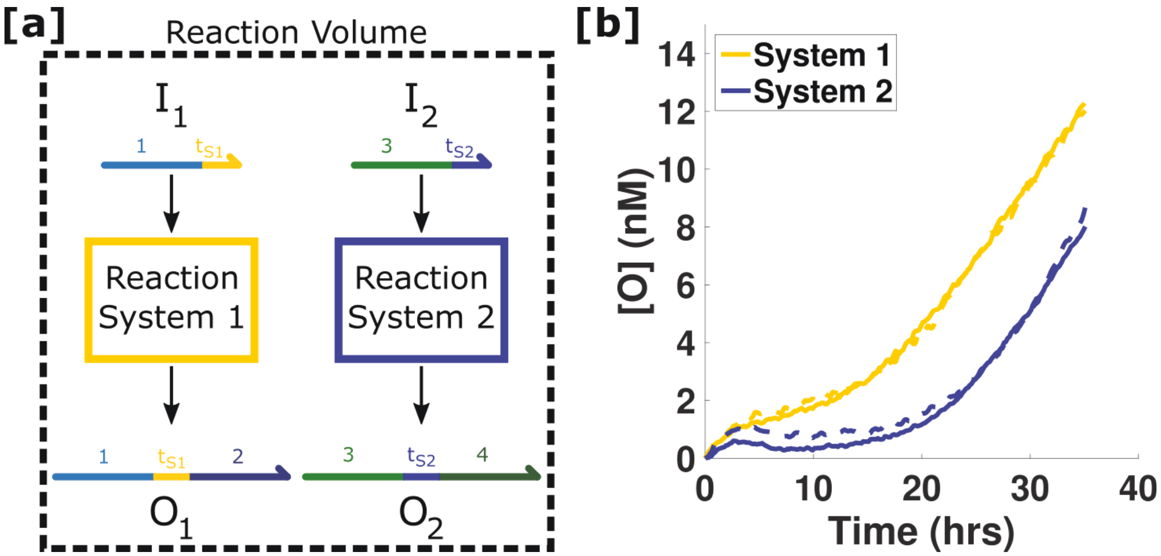


Figure 3.4: Multiplexed timers. *(a)* Schematic of multiple timer circuits operating within the same solution, releasing independent output strands at different times. *(b)* Two timers release output with the same rate but at different times (9 and 17 hours). Here, $[\mathbf{S}]_0=[\mathbf{I}]_0=0.5$ µM and $\{[\mathbf{D}]_{Sys1}, [\mathbf{D}]_{Sys2}\}=\{46$ nM, 63 nM$\}$. Dashed lines indicate the same reactions except with each system in isolation.

The timer circuit developed here successfully releases target strands of DNA into solution at a constant rate after a delay period. The sequence, delay time and production rate were easily tuned without needing to redesign the release system. We demonstrated delay times on the order of hours to days and production rates from a few nM/day to a hundred nM/day, which for volumes of 100 µL are approximately 0.1–10 nmol/day.

While the release of **O** observed (Fig. 3.3, Figures S3.8–S3.10) qualitatively followed our simple model described in Equations 3.1–3.5, there were differences between experiments and reaction curves predicted by the model. Notably we found that the release rate of **O** decreased faster and to a greater degree than the predicted reaction curves, suggesting the existence of uncharacterized reactions (*i.e.*, leak reactions) between **S** and **D** or the reporter. On the basis of the experiments, we hypothesize that DNA synthesis errors (*e.g.*, base-mismatches, truncations, additions or deletions) in the bottom strand of the **S** complex led to leak reactions and pathways that explained most of the deviation from the simple model. By accounting for these and other more minor leak reactions, a model was developed that provided a close fit with the experimental observations (Section 3.4.6), indicating that models can be used to program the rate and timing of output release.

3.3 Conclusion

By designing a circuit in which the output DNA sequence is an aptamer,[180,194]–*i.e.*, sequences of DNA that bind specifically to non-DNA species such as proteins[195] and other small molecules,[196] the circuit described here could also be used to control the dynamics of a wide range of other chemical systems beyond DNA strand-displacement.

Timer devices that can be programmed to release a particular species with a prespecified delay could be important for designing cascades for therapeutics or for self-assembly[90,91] in which different species are activated at different times. The constant low-rate of production we have shown could be used to design therapeutic hydrogels with novel, linear release profiles or within a reliable pulsed delivery system by combining it with a threshold amplifier system such as those used for signal restoration in molecular logic circuits.[87,114]

3.4 Supplemental Information

3.4.0: Materials and methods

The timer circuit was designed following previously outlined principles.[87] Sequences for each domain were drawn from Table S1 of the Supporting Online Material for Qian and Winfree, 2011, and are listed here in Table S1. Domains *S6* and *S5* listed below correspond to Domains *1* and *2* depicted in Figure 3.2 and Figure S3.0. The toehold (*t*) domain of System 2 was designed to have minimal non-specific interactions with the sequences of both systems using NUPACK.[116] Domain Names and Sequences are listed 5' to 3'. Additional schematics of the DNA complexes and the reactions are shown in SI 2.

Table S3.1: Sequence Data

Strand Names	Domain Names	Sequences
Source 1 Top	*S6 t S5*	CA TAACACAATCA CA TCT CA CCACCAAACTT CA
Source 1 Bottom	*t' S6'*	TG AGA TG TGATTGTGTTA TG
Initiator 1	*S6 t*	CA TAACACAATCA CA TCT CA
Delay 1 Top	*S5*	CA CCACCAAACTT CA CT
Delay 1 Bottom	*S5' t' S6' (2nt)*	AG TG AAGTTTGGTGG TG AGA TG TG
Reporter 1 Top	*S5 t Quencher*	CA CCACCAAACTT CA TCT CA/3IABkFQ/
Reporter 1 Bottom	*FAM t' S5' t'*	/56-FAM/TG AGA TG AAGTTTGGTGG TG AGA TG
Reporter 1 Full Complement	*t S5 t*	CA TCT CA CCACCAAACTT CA TCT CA
Source 2 Top	*S27 t S28*	AC AACACTCTATT AC AAT AC TCTACAATTCA AC
Source 2 Bottom	*t' S27'*	GT ATT GT AATAGAGTGTT GT
Initiator 2	*S27 t*	AC AACACTCTATT AC AAT AC
Delay 2 Top	*S28*	AC TCTACAATTCA AC CA
Delay 2 Bottom	*S28' t' S27' (2nt)*	TG GT TGAATTGTAGA GT ATT GT AA
Reporter 2 Top	*S28 t Quencher*	AC TCTACAATTCA AC AAT AC/3IABRQSp/
Reporter 2 Bottom	*TexasRed t' S28' t'*	/5TexRd-XN/GT ATT GT TGAATTGTAGA GT ATT GT
Reporter 2 Full Complement	*t S28 t*	AC AAT AC TCTACAATTCA AC AAT AC

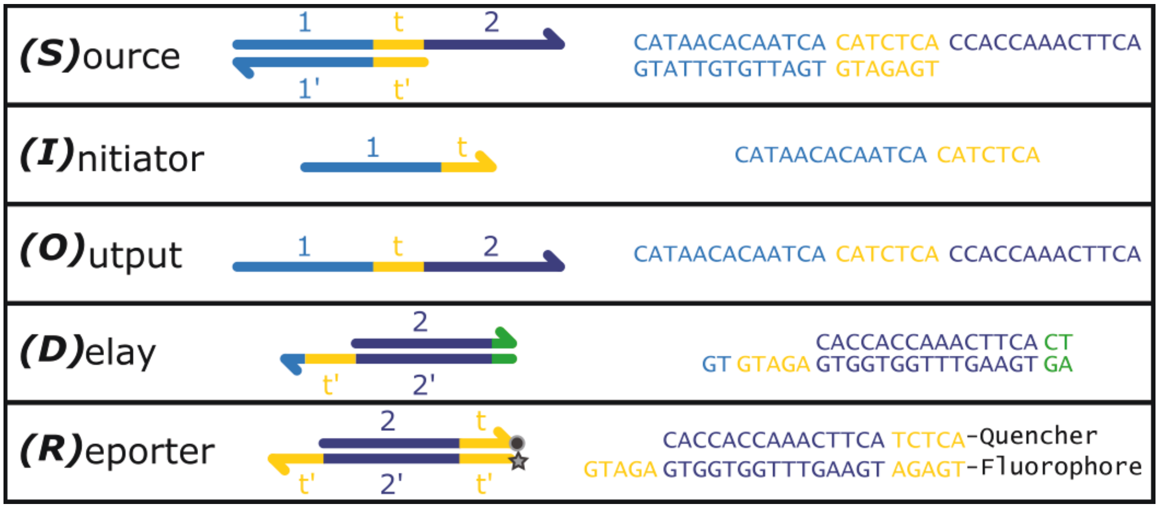


Figure S3.0: Schematic of reaction species with their sequences as listed in Table S3.1. Reactions between species for the timer circuit are shown in Figures S3.3 and S3.6.

Sequences were ordered as lyophilized powder from Integrated DNA Technologies (IDT). The Reporter Top and Reporter Bottom strands were ordered purified by high-performance liquid chromatography (HPLC) and all other strands were ordered impure with standard desalting. Strands were suspended in Millipore purified water to a concentration of ~1 mM and stored at -20 °C. Empirical oligonucleotide stock concentrations were determined by assaying the absorbance (OD260) of 1000x diluted samples of each stock solution at 260 nm. The extinction coefficient provided by IDT was used to calculate stock concentrations using the Beer-Lambert law.

Source, Delay and Reporter complexes were prepared at a concentration of 100 µM in Tris-acetate-EDTA buffer with 12.5 mM Mg^{2+} (1x TAE/Mg^{2+}). Each complex was annealed in an Eppendorf Mastercycler PCR by holding the solutions at 90 °C for 5 minutes followed by cooling at -1 °C per minute down to 20 °C. After annealing, the Source complexes were incubated with 100 µM of their complementary Initiator strand overnight at room temperature to react with any poorly formed Source complexes. The Source complex was not incubated with any other strands or complexes due to the increased complexity of gel purification.

After annealing each complex and incubating the Source complexes overnight, all complexes were purified by polyacrylamide gel electrophoreses (PAGE). Ten percent polyacrylamide gels were cast by mixing 3.25 mL of 19:1 40% acrylamide/bis solution (Bio-Rad) with 1.3 mL 10x TAE/Mg^{2+} and 8.45 mL Millipore-purified H_2O. This solution was polymerized by the addition of 78 μL 10% ammonium persulfate (APS) and 5.4 μL tetramethylenediamine (TEMED) in a gel cassette with a large single well comb at the top of each gel. Fifteen percent polyacrylamide gels were prepared in a similar fashion except with a corresponding higher fractional volume of 40% acrylamide/bis stock solution. Two hundred microliters of annealed DNA complexes were mixed with 6x loading dye (New England Biolabs, product #B7021S) and loaded into the wells of the gels in a Scie Plas TV100K cooled vertical electrophoresis chamber. The gels were run at 150V and 4 °C for 1.5 or 3 hours for 10% and 15% polyacrylamide gels. Reporter and Delay complexes were purified using 10% gels and the Source complex was purified using 15% gels. After running for the appropriate time, the bands were cut out using UV-shadowing at 254 nm for visualization. Bands were diced into ~1 mm^3 pieces, mixed with 500 μL of 1x TAE/Mg^{2+} buffer and were shaken on a vortexer overnight at room temperature. The DNA solutions were then transferred by pipet to a fresh tube leaving behind the gel pieces. The solutions were centrifuged for 5 minutes at 3000x*g* to precipitate any remaining gel pieces. The DNA solutions were transferred to a new tube and stored at 4 °C until use. The concentrations of these purified complexes were then measured with an Eppendorf Biophotometer with a dilution factor of 30x using the approximate extinction coefficient (ε):

$$\varepsilon_{\text{Final}} = \varepsilon_{Top-strand} + \varepsilon_{Bottom-strand} - 3200N_{AT} - 2000N_{GC}$$

where *N* indicates the number of hybridized A-T or G-C pairs in each complex.[197]

Reaction kinetics were measured on quantitative PCR (qPCR) machines (Agilent Stratagene Mx3000 and Mx3005 series) at 25 °C. Reactions were prepared in 96-well plates using 150 μL/well volumes. Each well contained 1x TAE/Mg^{2+} and 2 μM $PolyT_{20}$ strands to help displace reactant species from the pipet tips used to add species and potentially from the well walls. In a typical experiment, Millipore-purified H_2O, TAE/Mg^{2+} and $PolyT_{20}$ strands were first mixed together. Reporter complexes were then added at 100 nM for System 1 or 200 nM for System 2. Baseline fluorescent measurements of the Reporter complex alone was conducted for each experiment for 0.5 to 1 hour with measurements every 1 to 10 minutes. This baseline was taken to be where the [Output] is equal to zero as detected by the Reporter and was subtracted from all subsequent data. After measuring this baseline, DNA strands or complexes were added to each well, depending on the experiment (see Sections 3.4.1–4). Fluorescence measurements were taken every 1–5 minutes for Delay characterization or every 5–10 minutes for Production characterization and Timer experiments.

3.4.1 Reporter calibration

The Reporter complex (Figure 3.2c) was used to indirectly measure the concentration of the single-stranded Output strands in solution as a function of time. A Reporter that reacts reversibly with the Output species was used as an irreversible reporter could compete with the Delay species. The reporter follows the reaction:

$$Reporter + Output \underset{k_\mathrm{r}}{\overset{k_\mathrm{f}}{\rightleftharpoons}} \text{Fluorophore} + \text{Quencher} \tag{S3.1}$$

where the quencher-modified top strand of the Reporter complex is displaced by an invading strand causing in increase of fluorescence. The forward reaction rate constant, k_f, is expected to be around $5\mathrm{x}10^4$ $M^{-1}s^{-1}$.[96] Two calibrations were conducted to translate measured fluorescence intensities to levels of free Output concentration for each experiment. It was assumed that the measured fluorescence was proportional to the concentration of unquenched fluorophore, [Fluoro.] (*e.g.*, *Fluorophore* in equation S3.1), through a proportionality constant α. To determine α, we measured the fluorescence of the Reporter complex with known concentrations of the full complement (FC) to the bottom strand of the Reporter (see Table S3.1) and measured the change in fluorescence before and after addition of the complementary strand (Figure S3.1). In general, we used the equation

$$[FC] = \alpha * \Delta Fluorescence + \beta \qquad (S3.2)$$

for [FC] equal to 0, 25, 50, 75 and 100 nM. In the ideal case, β is equal to zero. Alpha was determined by calculating the slope after fitting a line to [FC] *vs.* ΔFluorescence (Figure S3.1). This calibration enables the normalization of all fluorescence data into [Fluoro.]. Additionally, Figure S3.1 shows that photobleaching of the fluorophores are not a significant factor in measuring fluorescence as seen by the stable, non-decreasing intensity values.

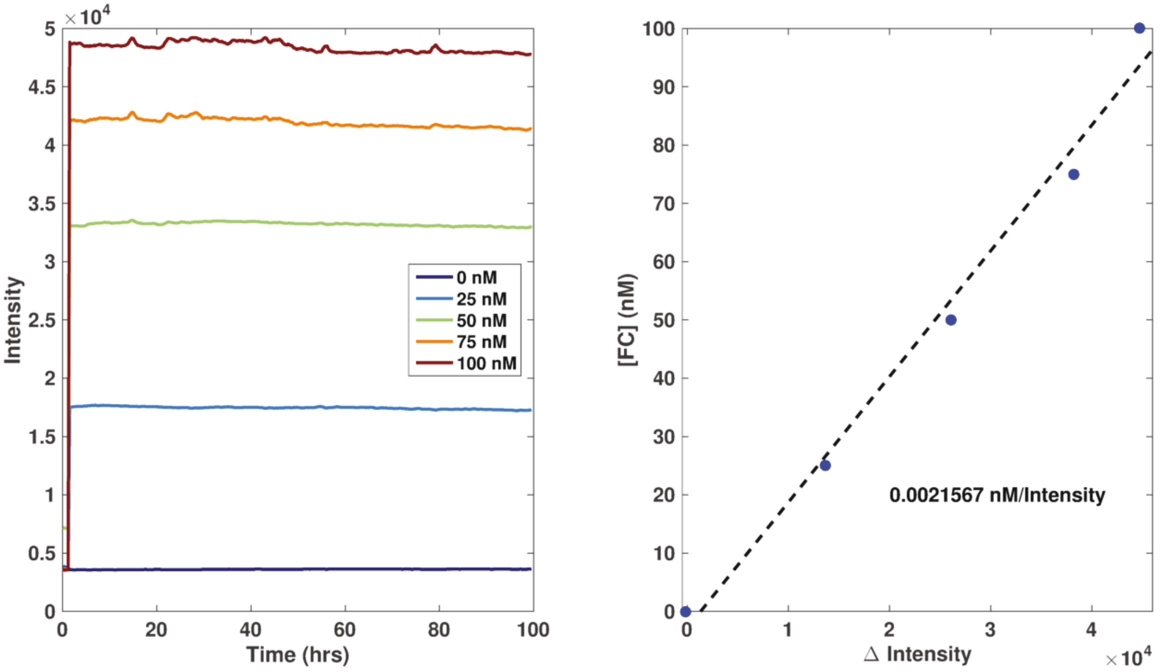


Figure S3.1: Example calibration plots for the Reporter complex with its full complement to convert raw intensity values into [Fluoro.]. FC was added to 100 nM Reporter at concentrations ranging from 0–100 nM as noted in the legend.

To convert the [Fluoro.] into [Output], the K_{eq} for the reporter reaction shown in equation S3.1 was calculated by mixing the Reporter complex with known concentrations of Output strand and using the equation:

$$K_{eq} = \frac{[Fluoro.]^2}{([O]_0 - [Fluoro.])([R]_0 - [Fluoro.])} \qquad (S3.3)$$

where $[\boldsymbol{O}]_0$ is the concentration of Output (*e.g.*, 25, 50, 75 or 100 nM) added to the Reporter solution and $[\boldsymbol{R}]_0$ is the initial Reporter complex concentration (*e.g.*, 100 nM). In general, we found that the intensities measured with this calibration method decreased over the duration of the experiment (Fig. S3.2a), possibly due to Reporter complexes becoming stuck in the "off" state. Due to this decrease, the K_{eq} was calculated as a function of time. The data was segmented into 75 bins (~1.3 hours each), with each bin having a K_{eq} calculated as the average K_{eq} over the [Output] tested (Fig. S3.2b,c).

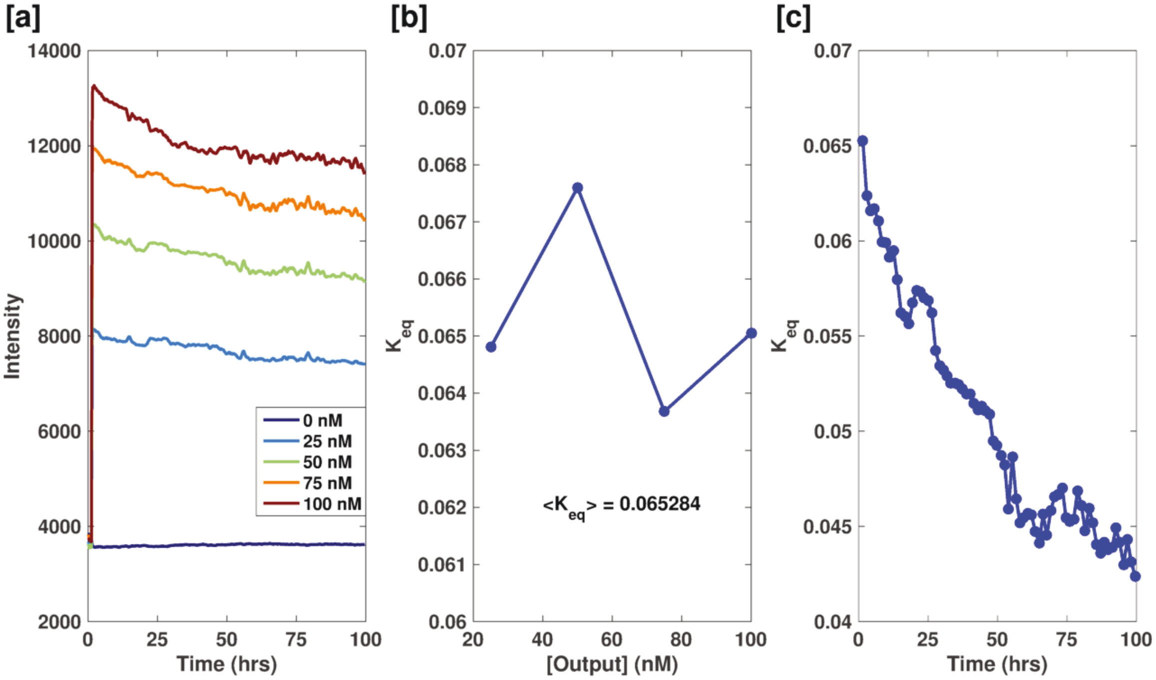


Figure S3.2: Example calibration plots for the Reporter complex with Output to convert [Fluoro.] values into [Output]. (a) Output was added to 100 nM Reporter at concentrations ranging from 0–100 nM as noted in the legend. (b) For each time segment, the K_{eq} was calculated as the average value over the [Output]'s tested. Little variance was seen between the K_{eq}'s calculated at each [**O**] within a given time segment. The initial segment is shown (t = 1 hr). (c) The K_{eq} decreases as a function of time. Each segment usually contained 1.3–1.5 hours of data.

The concentration of Fluoro. was then converted into [Output] through the equation

$$[O] = \frac{[Fluoro.]^2}{K_{eq}([R]_0 - [Fluoro.])} + [Fluoro.] \qquad (S3.4)$$

where *[Fluoro.]*, *[**O**]* and K_{eq} are functions of time. This equation reports the total concentration of Output as the sum of free **O** in solution and **O** that is transiently bound to the Reporter complex.

3.4.2 Production and delay reaction characterization

To understand the timer system, we initially characterized the production reaction (between the Source and Initiator molecules) without a concurrent delay reaction. Multiple production reactions using various concentrations of Initiator and Source were conducted to calculate the forward and reverse reaction rate constants for the set of reactions shown in Figure S3.3. Two sets of reactions were conducted: [**S**]=[**I**] and [**S**]x[**I**]=1. In the first case, the production rate (k_{prod}) varies because the product of the initial concentrations of Initiator and Source is changing. In the second case, the production rate is theoretically constant between experiments on "short" time scales.

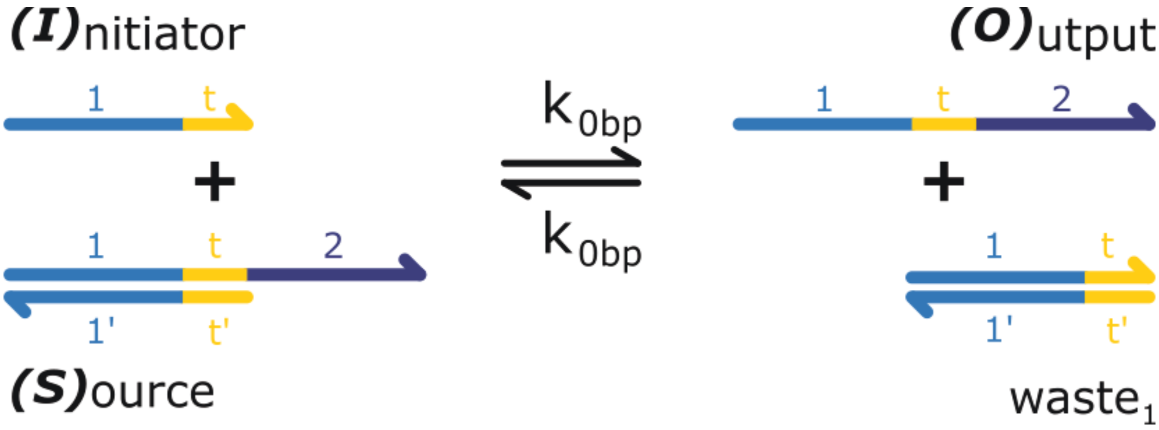


Figure S3.3: Schematic for the Production circuit. The Initiator reacts with the Source complex through a fraying mechanism at the ends of the double-stranded regions to produce Output and waste$_1$. The forward and reverse reaction rate constant depend on DNA sequence and the point of strand-displacement initialization and thus could be different values, but for simplicity a single rate constant, k_{0bp}, was chosen.

The Source and Reporter were incubated for about 6 hours until the measured intensity reached a steady state prior to the addition of **I**. This steady state intensity is thought to be a small population of free Output left over from the purification process, although interactions between the Source and Reporter complexes may exist. The concentration of detected Output by the Reporter corresponded to 0.01-0.02x[Source] used in each experiment (Figure S3.4).

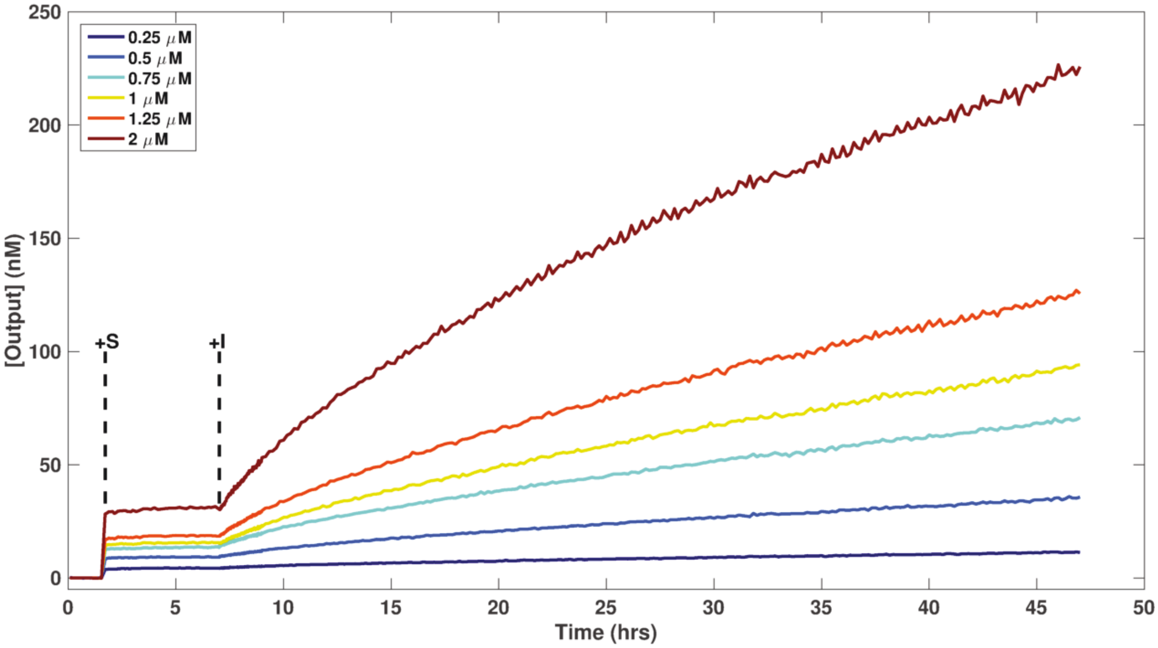


Figure S3.4: Characterization of the effect of [Source] and [Initiator] on the observed release of Output. The Source is mixed with the Reporter after 1.5 hours and incubated for about 6 hours, when the Initiator is added (denoted by black dashed lines). For the curves shown, [**S**] is equal to [**I**] and is shown in the legend. The concentration of initial Output detected prior to initiation scales with the amount of [**S**] added. Data is identical to that shown in Figure S3.5a and Figure 3.3a.

The post-initiation reaction curves were fit using MATLAB to initially calculate k_{0bp} for this set of reactions using the second-order reaction kinetics equation:

$$\frac{d[O]}{dt} = k_{0bp}[I][S] - k_{0bp}[O][waste_1] \qquad (S3.5)$$

The Reporter reaction was not included in the fitting procedure because its equilibration kinetics were assumed to be much faster (~10^5x faster reaction rate constant) than that of the Production reaction. However, the reaction rate constants calculated from this model did not capture the dynamics seen in Fig S3.5a–b–*i.e.*, a quick release of **O** followed by a slower, more linear region. We hypothesized that this was due to a small concentration of Source that reacted quickly and irreversibly with the Initiator present in the reaction volumes, perhaps because of errors in sequence produced during solid state DNA synthesis. We call this small population [$\mathbf{S_{Leak}}$] which reacts with reaction rate constant k_{Leak}. To account for this possibility, we used the following equation to generate a better fit to the experimental data:

$$\frac{d[O]}{dt} = k_{0bp}[I][S] - k_{0bp}[O][waste_1] + k_{Leak}[S_{Leak}][I] \qquad (S3.6)$$

and to calculate the model parameters k_{0bp}, k_{Leak} and [$\mathbf{S_{Leak}}$]. Additional information regarding $\mathbf{S_{Leak}}$ can be found in Section 3.4.6. An example of such a fit is shown in Figure S3.5c and the fit parameters are compiled in Table 2. The average k_{0bp} was 0.49±0.13 1/M-sec which is in good agreement with the reported value of 0.5 1/M-sec.[96] [$\mathbf{S_{Leak}}$]$_0$ varied from 0.5 to 4.7% of [$\mathbf{S}$]$_0$.

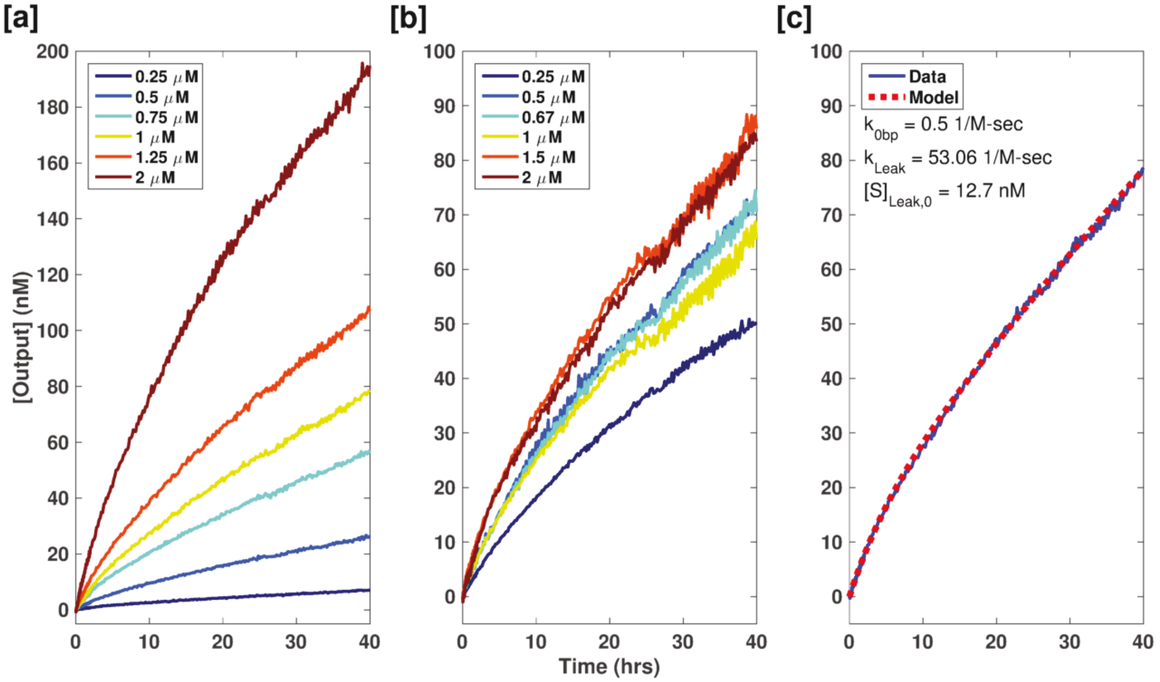


Figure S3.5: Characterization of Production reactions. Production reactions were measured for [Source] = [Initiator] in the range of 0.25–2 μM (a) and for [**S**] x [**I**] = 1 (b). The legend in (b) shows the concentration of Source in the reaction mixture. (c) Example comparison between experimental measurements and the model resulting from the fit of reaction rate constants for [**S**]=[**I**]=1 μM. Reaction curves were fit using a bimolecular reaction kinetics model to calculate a second-order reaction rate constant. Reaction rate constants for each curve are listed below in Table S3.2.

As shown in Figure S3.5a, the production rate varied with the concentrations of Source and Initiator. From equation 3.5, the production rate is expected to follow a power law ($k_{\mathbf{Prod}} \sim [\boldsymbol{S}]_0^2$) when initial concentrations of **S** and **I** are equal. The experimentally determined production rate increased to the exponential of 1.6 with increasing **S** and **I** instead of 2. There was also a small variation in the production rate for the case [**S**]x[**I**]=1, where the initial k_{prod} was expected to be constant among the reaction conditions tested (see Table S3.2). The deviation from the expected result in both cases could be due to other undesired reactions present in between reaction species.

Table S3.2: Reaction rate constants for the Production reactions shown in Figure S3.5. The production rate constant (k_{prod}) was calculated using equation 3.5 using the initial concentrations of Source and Initiator. Values listed are result of fit with 95% confidence interval bounds. Overall values are the average and standard deviation of all rows (k_{0bp}) or rows in the [**S**]x[**I**]=1 set of experiments (k_{prod}).

Experiment Type	[**S**] (μM)	[**I**] (μM)	k_{0bp} (1/M-sec)	k_{prod} (nM/hr)	k_{Leak} (1/M-sec)	[$\mathbf{S_{Leak}}$] (nM)
[**S**]=[**I**]	0.25	0.25	0.676±0.008	0.152±0.002	336±20	1.3±0.1
	0.5	0.5	0.645±0.006	0.58±0.005	129±7	4.6±0.2
	0.75	0.75	0.647±0.004	1.311±0.009	91±4	8.8±0.2
	1	1	0.501±0.004	1.803±0.014	53±2	12.7±0.4
	1.25	1.25	0.439±0.004	2.47±0.021	43±2	19.0±0.6
	2	2	0.279±0.005	4.016±0.066	20±1	53.8±2.0
[**S**]x[**I**]=1	0.25	4	0.32±0.01	1.14±0.02	8±1	11.7±0.8
	0.5	2	0.46±0.01	1.67±0.02	27±2	13.2±0.6
	0.67	1.5	0.46±0.01	1.65±0.02	35±3	12.6±0.6
	1	1	0.39±0.01	1.41±0.03	52±4	14.1±0.7
	1.5	0.67	0.53±0.01	1.91±0.03	96±6	18.2±0.8
	2	0.5	0.55±0.01	1.97±0.02	129±7	15.9±0.6
Overall			0.5±0.1	1.6±0.3		

The Delay circuit module was characterized by the degree to which the Delay complex was able to sequester free Output. The kinetics of the reaction between Delay and Output was too fast to capture using the concentration ranges tested in order to fit reaction rate constants for the set of reactions show in Figure S3.6. The Delay reaction is slightly reversible due to 2 extra bases on the Delay complex that are not complementary to the Output (shown in green). These bases are important for decreasing undesired reactions between the Source and Delay complexes, whose reaction rate constant is on the same order as the Production circuit (initialized *via* fraying). However, this reaction set is expected to be very forward reaction dominated since k_{forward} ~2x10^{6} 1/M-sec and k_{reverse} ~10^{2} 1/M-sec.[96]

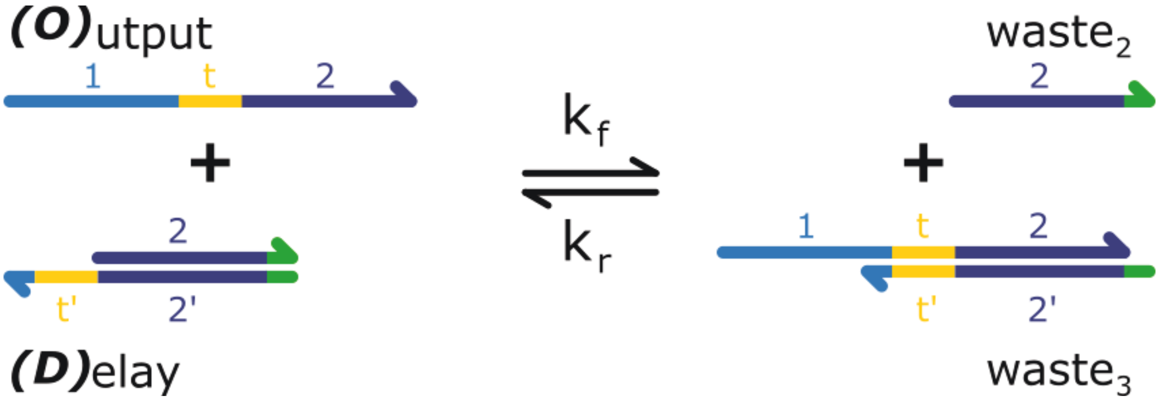


Figure S3.6: Schematic for the Delay circuit. The Output reacts with the Delay complex through a 7 base-pair toehold to produce two waste species. Two extra bases on the Delay complex (shown in green) inhibit a leak reaction between Source and Delay complexes.

Two sets of experiments were run to investigate whether the Delay complex efficiently sequestered Output in a reaction solution. The first set was run by adding **D** at various concentrations to the Reporter followed by the addition of **O**. As shown in Fig. S3.7a, the concentration of **O** detected closely matched the expected result: [**O**] = $[\mathbf{O}]_0$-[**D**]. The second set of experiments was run by first adding **O** to the Reporter followed by the addition of **D**, leading to a sudden decrease in fluorescence intensity and detected free **O** (Fig. S3.7b). Again, the remaining [**O**] matched the expected concentrations determined by the amount of **D** added. The concentration of **O** remaining when the **D** exceeded the initial concentration of **O** added did not fully decrease to zero due to the reversibility of the Delay reaction (Fig. S3.6).

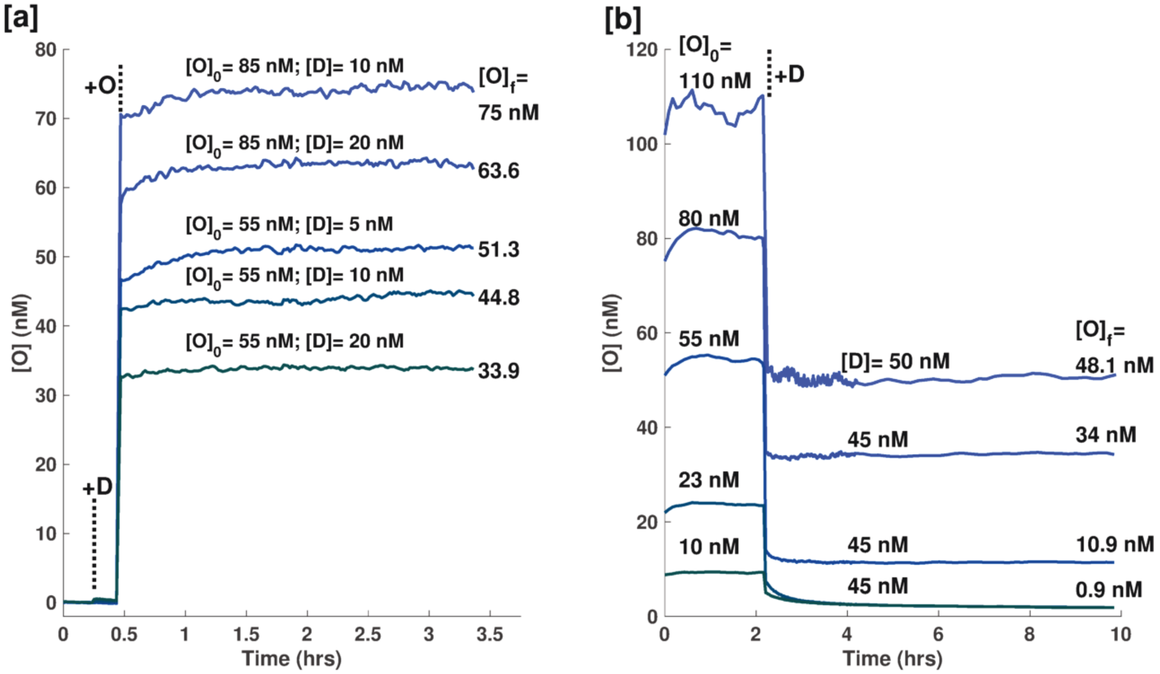


Figure S3.7: Characterization of the Delay circuit. (a) **O** is added to **R** and **D** leading to a rise in detected **O** by the Reporter. Black dashed lines indicate the times **D** or **O** is added to the Reporter solution. The concentration of **O** and **D** added to each reaction mixture is annotated above each trace. [**O**]$_f$ is the average [**O**] over the last 30 data points. (b) **D** is added to Reporter and **O** leading to a decrease in fluorescence as **O** is being sequestered by **D**. The data in Fig. 3.3b is calculated from the data shown here in (a) and (b).

3.4.3 Timer experiments with system 1

Timer experiments were conducted similar to Production experiments, except both Source and Delay were mixed with the Reporter after the initial baseline was measured (~1 hour). Initiator was mixed into the wells after the intensity reached a steady-state, at about 22–24 hours. Both **S** and **I** were kept equal in these experiments. From Table S3.2, [**S**]=[**I**]=1 μM produces Output at ~1.8 nM/hr, [**S**]=[**I**]=0.5 μM at ~0.6 nM/hr and [**S**]=[**I**]=0.25 μM at ~0.15 nM/hr (Fig. 3.3c-d and Figs. S3.8–10).

The delay time (t_{Delay}) was determined by calculating a moving linear fit of each curve post-Initiator addition. The span of each fit included 100 or 60 data points, depending on the reaction conditions. The step size was ½ of the span in each case (50 and 30 points). The xintercept from the fit with the largest slope was chosen as the delay time. Other algorithms (*e.g.*, the time [**O**] or d[**O**]/dt surpasses a specified value) gave similar time delays, but were more sensitive to noise/bias.

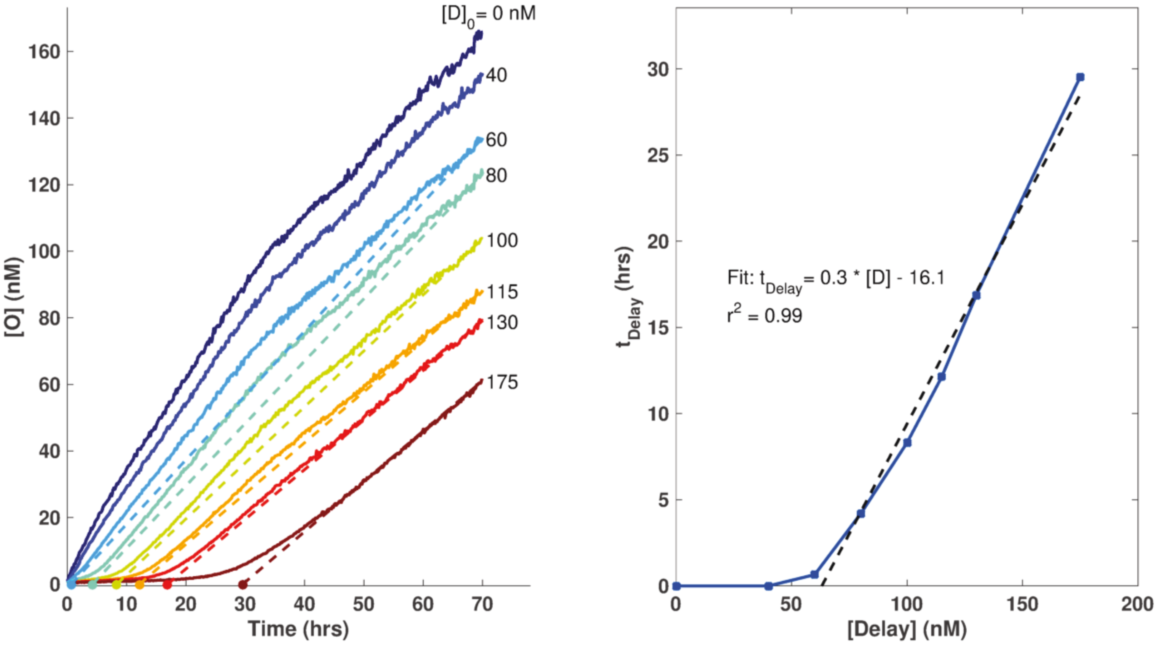


Figure S3.8: [Output] *vs.* time and t_{delay} *vs.* [Delay] for reactions using [**S**]=[**I**]=1 μM production conditions for System 1. Dashed lines in the left plot are guides showing the calculated time delay. Dashed line in the right plot shows a linear fit for points with t_{Delay} greater than zero. Data is the same as in Figure 3.3c,d.

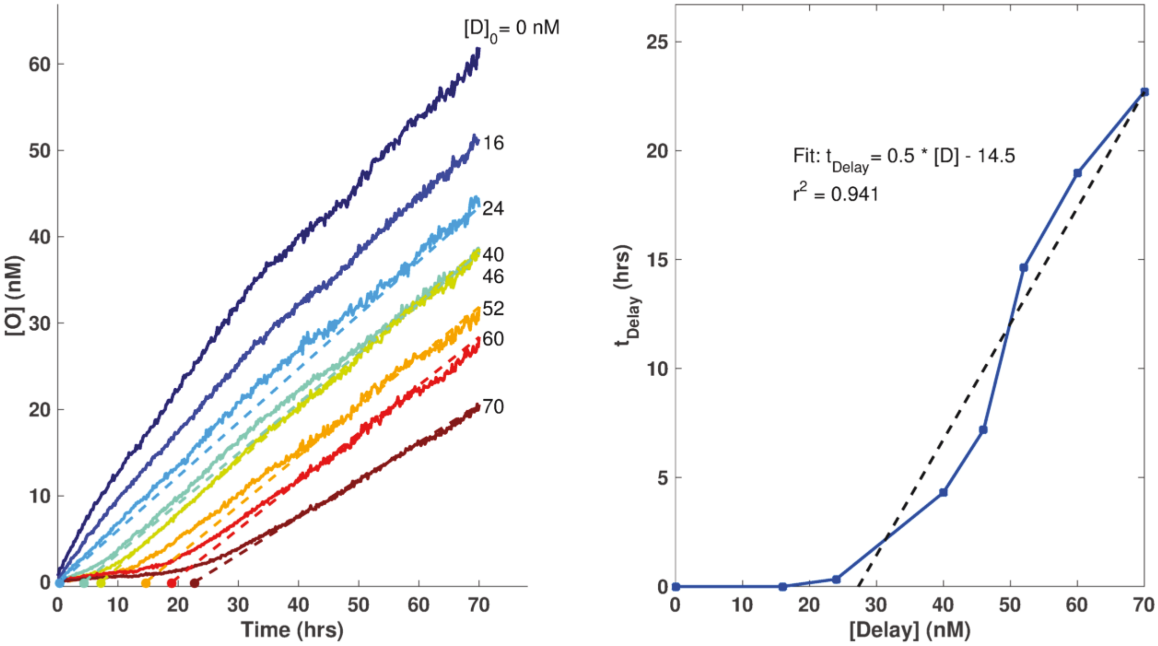


Figure S3.9: [Output] *vs.* time and t_{delay} *vs.* [Delay] for reactions using [**S**]=[**I**]=0.5 μM production conditions for System 1. Dashed lines in the left plot are guides showing the calculated time delay. Dashed line in the right plot shows a linear fit for points with t_{Delay} greater than zero.

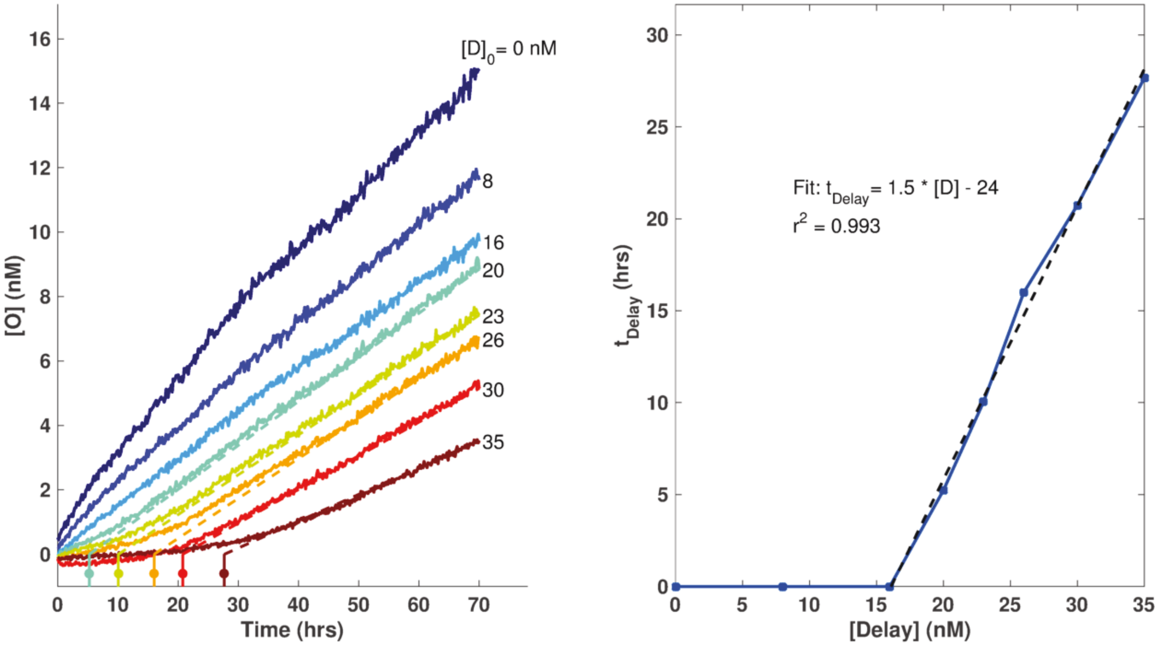


Figure S3.10: [Output] *vs.* time and t_{delay} *vs.* [Delay] for reactions using [**S**]=[**I**]=0.25 μM production conditions for System 1. Dashed lines in the left plot are guides showing the calculated time delay. Dashed line in the right plot shows a linear fit for points with t_{Delay} greater than zero.

The rate of production (d[**O**]/dt) was calculated for each reaction condition. The slope of each curve was calculated as the average over the last 90 data points (15 hours). We found that the production rate at that time point decreased slightly as a function of [Delay], possibly due to the uncharacterized reactions mentioned elsewhere. Despite the Delay dependent effects observed, an appropriate [Delay] can be chosen from a desired production rate and time delay using Table S3.2 and Figure S3.11.

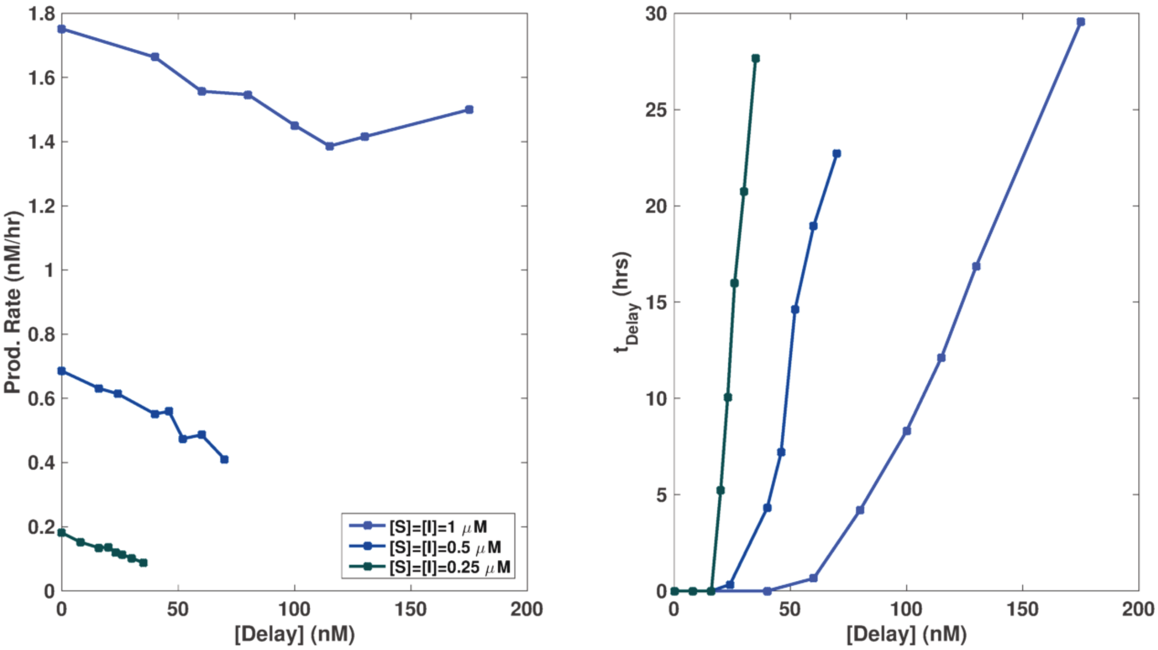


Figure S3.11: Production rate *vs.* [Delay] and t_{delay} *vs.* [Delay] for System 1. Production rate was calculated as the average over the last 90 data points (15 hours). Production rate decreases with [Delay] due to possible undesired reactions between circuit components.

3.4.4 Timer experiments with system 2

Experiments for System 2 were conducted the same as with System 1 except with a [Reporter] of 200 nM. Output production from System 2 was found to be slower than with System 1 (Figures S3.12–14). This could be due to DNA sequence differences between the systems; notably the toehold domain of System 2 is expected to have weaker binding than that of System 1 because it has less G-C base content. The 7 bp toehold of the Delay complex is weaker as well. Additionally, if there are significant interactions between the Source or Delay and the Reporter complex, the increased Reporter concentration could be an attributing factor. Finally, while System 1 had a decrease in production rate with increasing [Delay], System 2 showed an increase in production rate (Figure S3.15).

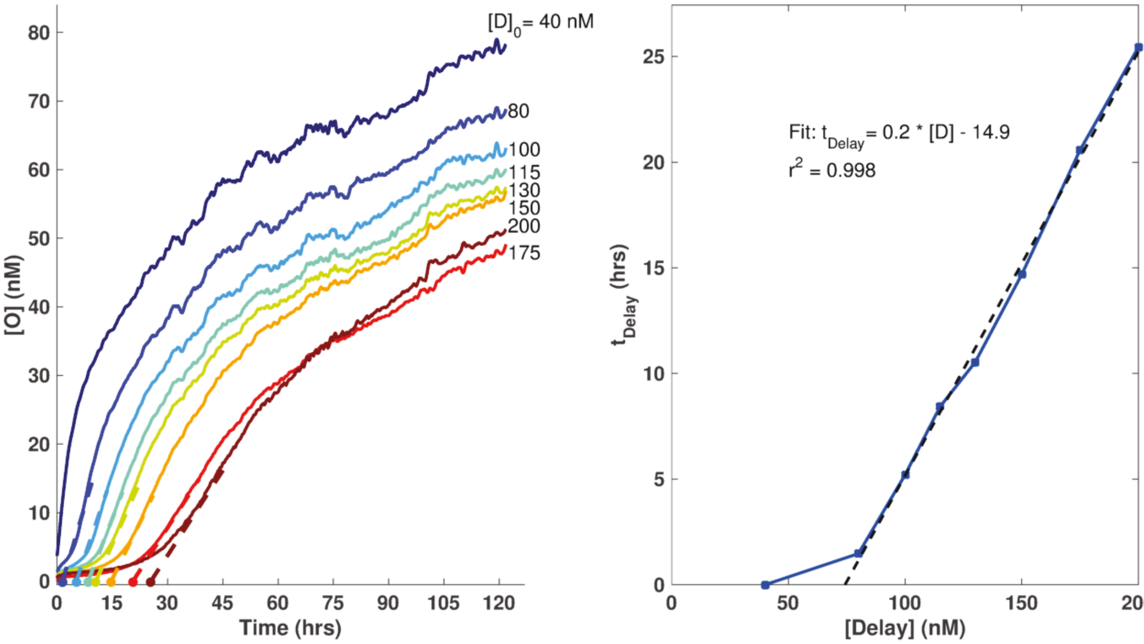


Figure S3.12: [Output] *vs.* time and t_{delay} *vs.* [Delay] for reactions using [**S**]=[**I**]=1 μM production conditions for System 2. Dashed lines in the left plot are guides showing the calculated time delay. Dashed line in the right plot shows a linear fit for points with t_{Delay} greater than zero.

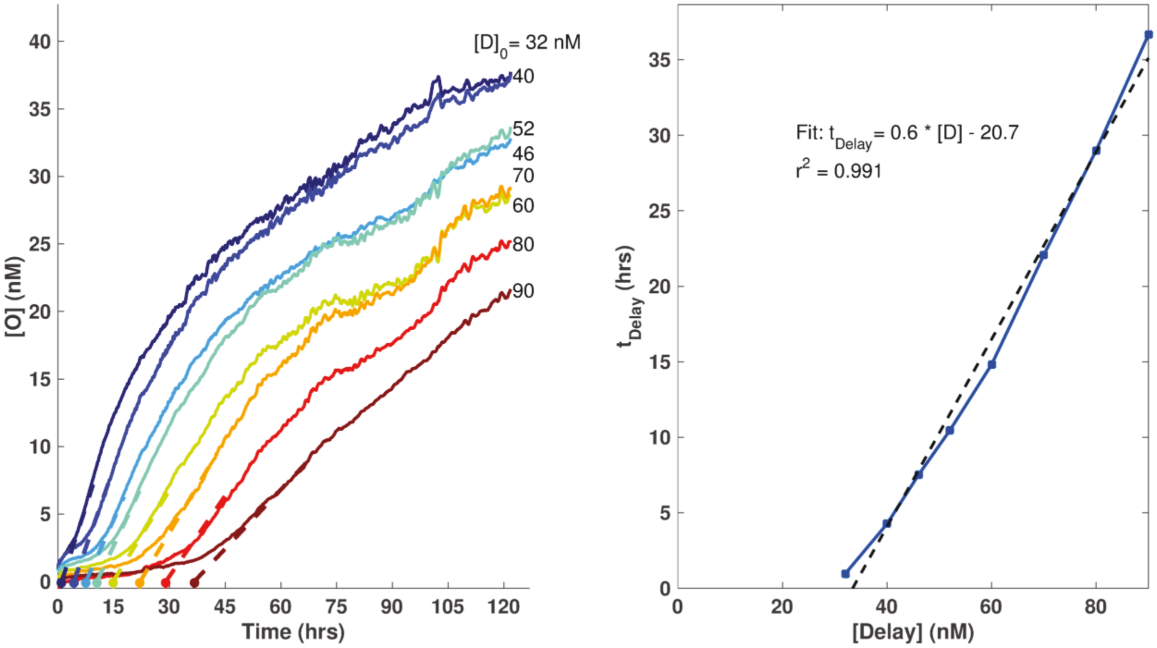


Figure S3.13: [Output] *vs.* time and t_{delay} *vs.* [Delay] for reactions using [**S**]=[**I**]=0.5 μM production conditions for System 2. Dashed lines in the left plot are guides showing the calculated time delay. Dashed line in the right plot shows a linear fit for points with t_{Delay} greater than zero.

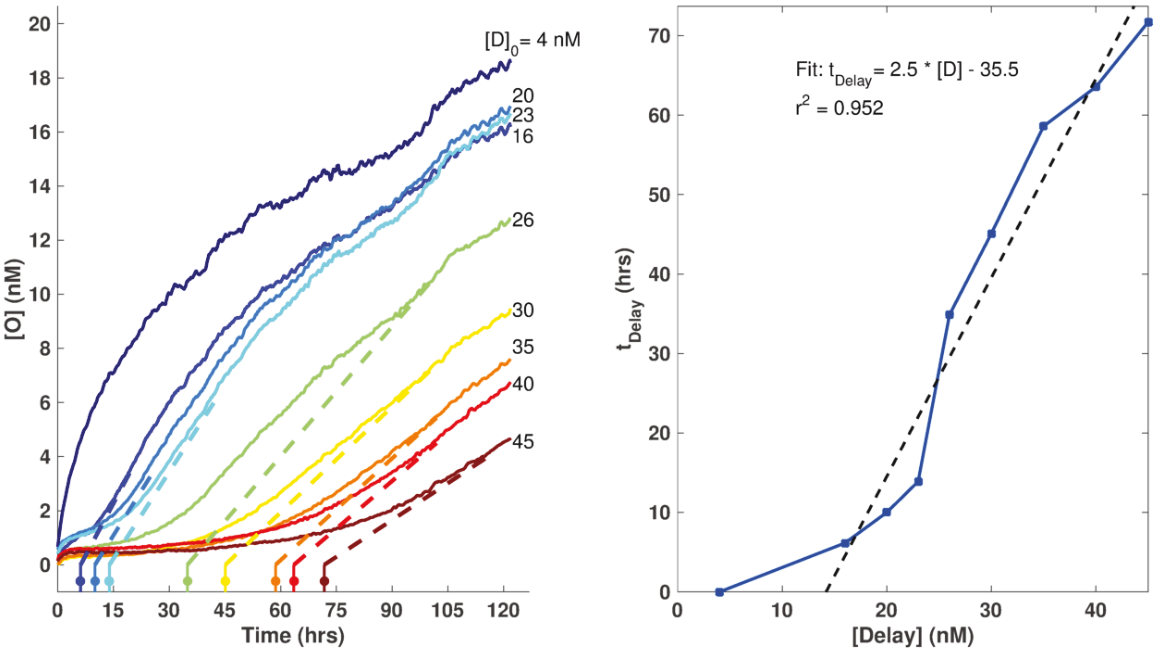


Figure S3.14: [Output] *vs.* time and t_{delay} *vs.* [Delay] for reactions using [**S**]=[**I**]=0.25 μM production conditions for System 2. Dashed lines in the left plot are guides showing the calculated time delay. Dashed line in the right plot shows a linear fit for points with t_{Delay} greater than zero.

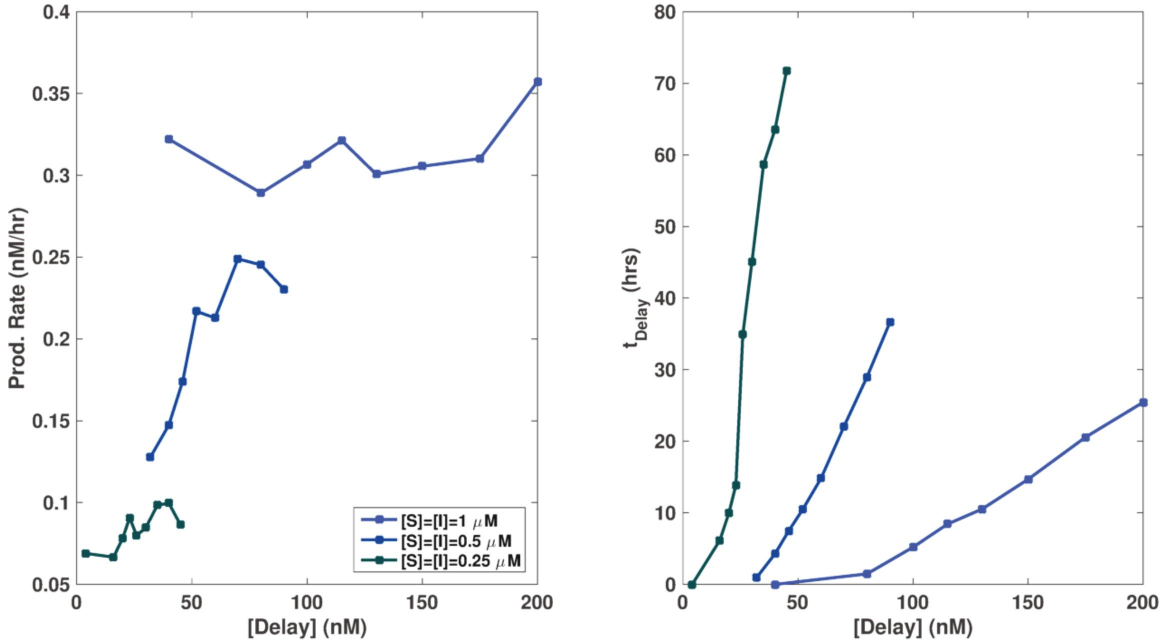


Figure S3.15: Production rate *vs.* [Delay] and t_{delay} *vs.* [Delay] for System 2. Production rate was calculated as the average over the last 250 or 100 data points (about 40 or 16 hours). Production rate increases with [Delay] due to possible undesired reactions between circuit components.

3.4.5 Multiplex timer experiments (systems 1 and 2)

Multiplexing experiments were conducted the same as the experiments described in Sections 3.4.3–4. Briefly, the Reporters of each system were mixed and a baseline was taken followed by the addition of Delay and Source complexes. Initiator was added after 22–24 hours. For experiments comparing data of each system in isolation *vs.* together, the reaction solution contained Reporters from both systems, but only the **D**, **S** and **I** from the system being studied (dashed lines in Fig. S3.16). System 1 was tracked using FAM and System 2 with TexasRed fluorophores using two different filters on the qPCR. Fluorescence from one fluorophore was not observed when measuring the fluorescence of the other fluorophore.

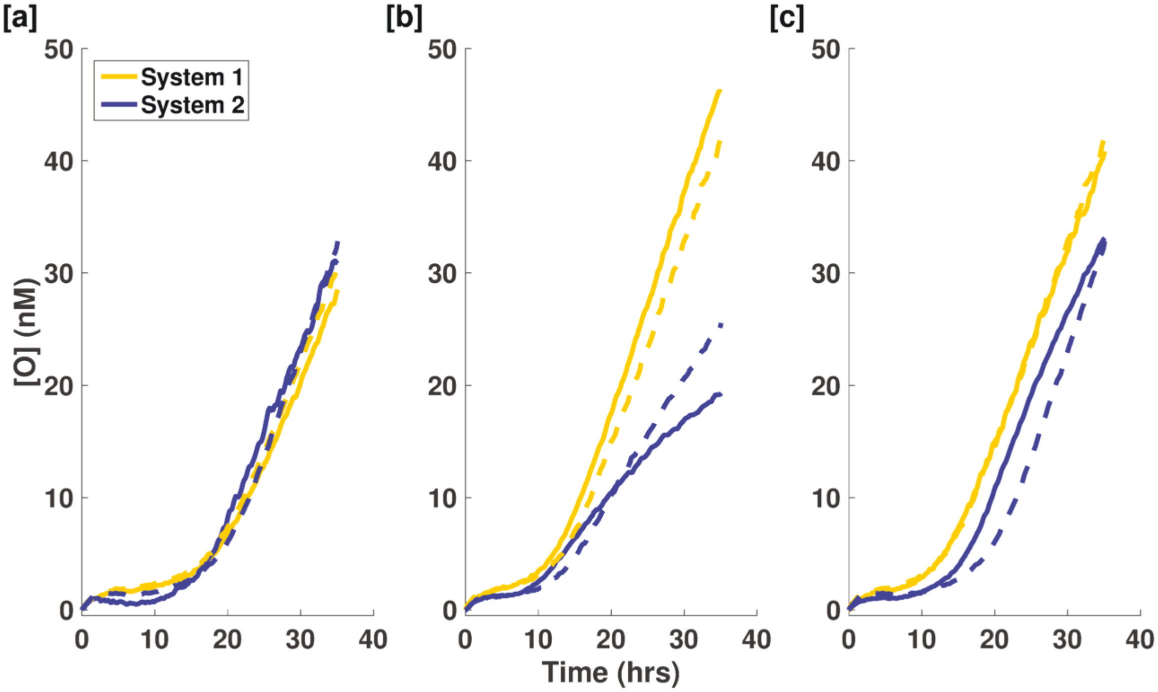
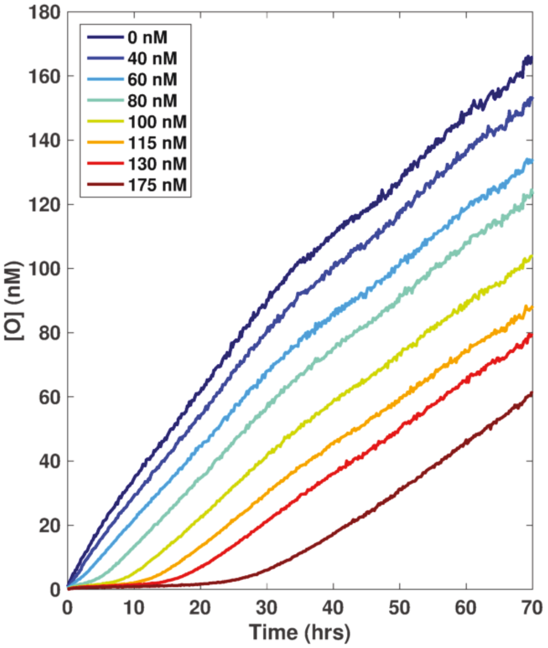


Figure S3.16: Additional examples of multiplexing two timer circuits. In each case, dashed lines indicate a reaction with the system in isolation. (a) Reaction mixture conditions were chosen such that both systems would release their respective **O** at the same time (19 hours) and rate. $[\mathbf{S}]_0=[\mathbf{I}]_0=1$ μM; $\{[\mathbf{D}]_{\mathrm{Sys1}}, [\mathbf{D}]_{\mathrm{Sys2}}\}=\{130$ nM, 165 nM$\}$. (b) The release rate of each system can be independently controlled while keeping the delay time constant (9 hours). $[\mathbf{S}]_0=[\mathbf{I}]_0=1$ μM for Sys1 and 0.5 μM for Sys2; $\{[\mathbf{D}]_{\mathrm{Sys1}}, [\mathbf{D}]_{\mathrm{Sys2}}\}=\{100$ nM, 50 nM$\}$. (c) The time of release (9 and 19 hours) of each system can be independently controlled while keeping the release rate constant. $[\mathbf{S}]_0=[\mathbf{I}]_0=1$ μM; $\{[\mathbf{D}]_{\mathrm{Sys1}}, [\mathbf{D}]_{\mathrm{Sys2}}\}=\{100$ nM, 165 nM$\}$.

3.4.6 Timer circuit simulations and characterization of leak reactions

Since the DNA strand-displacement circuit can be represented by a series of mathematical equations (equations S3.1 and S3.5 and Figure S3.6), a model that matches the experimental behavior of the system could be built to further tune the circuit for future applications. However, we found that a simple model derived from those equations failed to quantitatively capture the delay times observed in experiments (Figure S3.17). Based on that mismatch and the observation of an increase in fluorescence (or detected **O**) by the Reporter when **S** is added in the absence of **I** or **D** (Figure S3.4), we postulated that a series of interactions might exist between **S**, **D**, **I** and Reporter beyond what is predicted by the simple model.



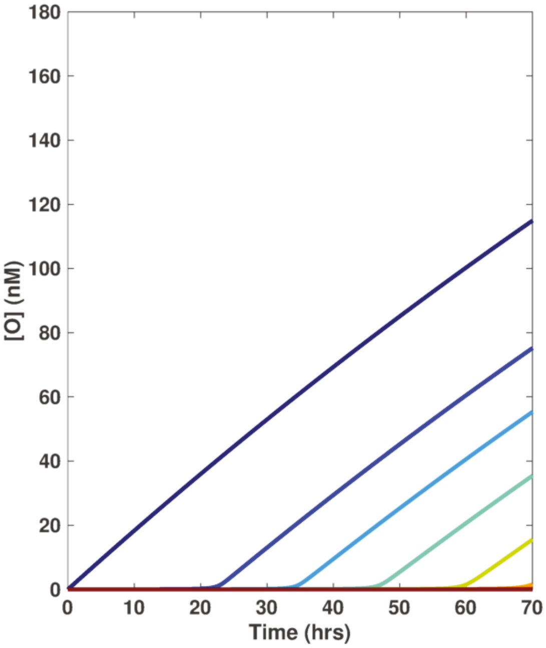


Figure S3.17: Comparison of data (left) and a model considering only abstract reactions described in eqn. S3.5 and Fig. S3.6 (right) for System 1 using 1 μM **S** and **I**. Delay concentrations for both plots are shown in the legend. Parameters for the model were as described in Section 3.4.2 (Figs. S3.3, S3.6) and in Zhang and Winfree.[96] Note the decrease in overall production and increase in time delay of the model prediction compared to experimental results.

In control experiments, we found that mixing the Reporter and Source quickly produced an observable fluorescence signal in the absence of Initiator (Fig. S3.4), suggesting that the Source complex and the Reporter interacted in an undesired or "leak" reaction (Fig. S3.18a). While by design, these species could interact through a 4-way 0 bp pair branch migration interaction, such a pathway would not explain the fast rate of reaction we observed. We therefore postulated that this reaction could be due to truncations or base mismatches within the toehold region of the bottom strand of some **S** complexes. We designated **S** complexes with these variations as the subspecies $\mathbf{S_{Leak}}$. These complexes would not have been separated from pure **S** complex during the purification process because their electrophoretic mobility is very similar to that of **S**.

We also observed that some leftover **O** remained after the gel purification process due to the proximity of the bands in the gel. Purifying **S** using a 15% polyacrylamide gel instead of a 10% gel significantly reduced the level of pre-initiation **O** detected, but some may still remain in the purified **S** complex solution. The leftover **O** and $\mathbf{S_{Leak}}$ complex are also expected to interact with the Delay complex for timer circuit reactions. We also considered a leak reaction between the Initiator and the Delay complex since there are 7 complementary nucleotides for a transient hybridization. The schematic shown in Figures S3.18–20 shows the possible leak reactions considered here. Unless specified in the figure captions, reaction rate constants for these reactions were taken from Zhang and Winfree (2009) or fit using the bimolecular rate equation shown in Section 3.4.2 and Figure S3.5. This approach of choosing reaction rate constants supports a physical representation of the postulated reactions and provides consistency with studies of other DNA strand-displacement reactions.

While the reactions described above may account for the unintended reactions that occurred within the timer system, including these reactions in a model still predicted significantly different time delay values than what were observed experimentally. To account for the decreased time delay observed in experiments, we added a small leak reaction between pure Source complexes and Delay complexes. While such a reaction would be expected to occur with a rate constant smaller than k_{0bp}, a reaction rate constant of k_{1bp} was needed to account for the large decrease in delay time. As noted in Section 3.4.2, we found that a reaction between $\mathbf{S_{Leak}}$ and **I** produced a better fit to the production dynamics. However, we would not expect this reaction to occur since **S** and **I** are incubated prior to PAGE purification of the **S** complex, any **S** species that would quickly react with **I** would be removed. Figure S3.19 shows an example comparison between experimental data (System 1, [**S**]=[**I**]=1 μM) and the resulting model prediction. While these reactions are only a possible description of the interactions between the DNA species, they show that an understanding of the reaction behavior is possible through the incorporation of leak pathways. We found that incorporation of each of these leak pathways into our model, using previously published rates and the fitted parameters of [$\mathbf{S_{Leak}}$], $k_{S,Leak}$ and k_{0bp} as described in Section 3.4.2 (no additional fitting parameters were required) produced quantitative agreement between our model and the experimental results that we observed (Fig. S3.21). Thus, we expect that this model can be used to tune the system's performance, including release rates and delay times.

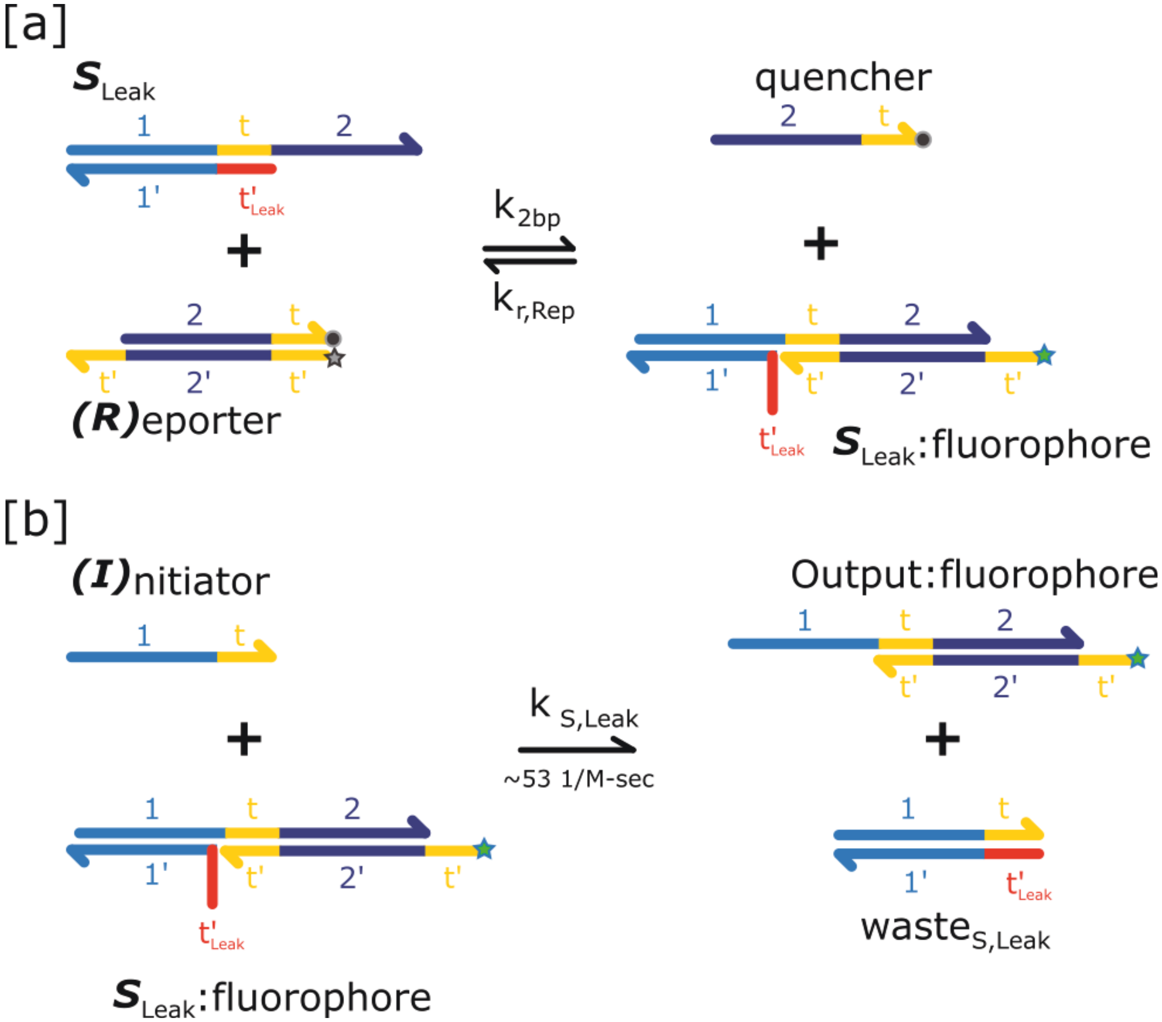


Figure S3.18: Reactions between a Source complex with mismatches or base truncations in the toehold region and the Reporter complex lead to the detection of a fluorescent signal prior to the addition of the Initiator. (a) $\mathbf{S_{Leak}}$ complex interacts reversibly with the Reporter to produce a fluorescent complex that reacts irreversibly with the Initiator (b). $k_{r,rep}$ was calculated from k_{5bp}[96] and the experimentally measured K_{eq} of the Reporter-Output reaction (Fig. S3.2).

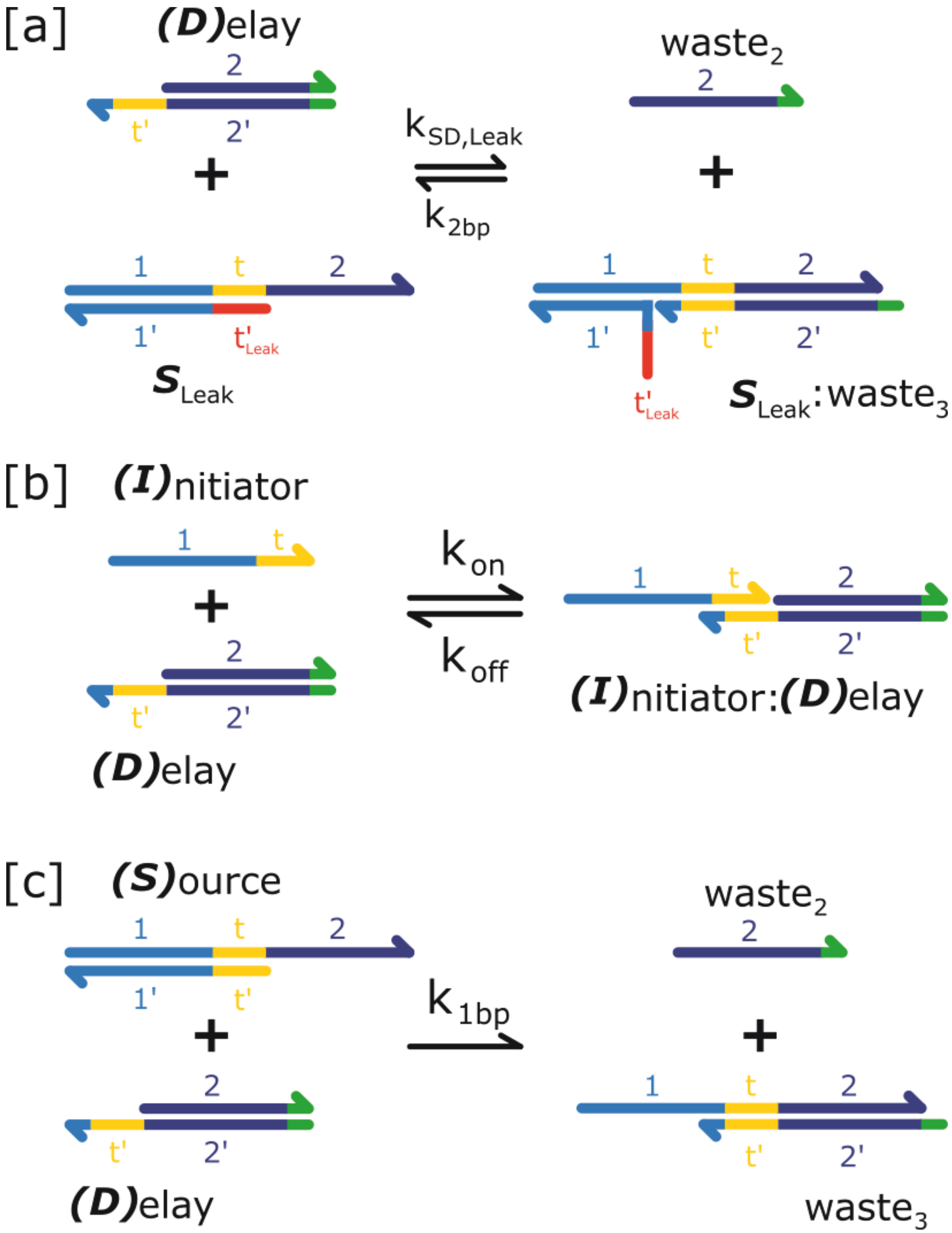


Figure S3.19: Reactions between the Delay complex and other reaction species. (a) Delay and $\mathbf{S}_{\mathbf{Leak}}$ react reversibly with a forward rate constant estimated to be $k_{SD,Leak}$ ~250 1/M-sec. (b) Initiator hybridizes and de-hybridizes with the toehold of the Delay complex. From Zhang and Winfree,[96] k_{on} ~ 3.5×10^6 1/M-sec and k_{off} ~ 0.08 1/sec. (c) Leak reaction between pure Source complex and the Delay complex.

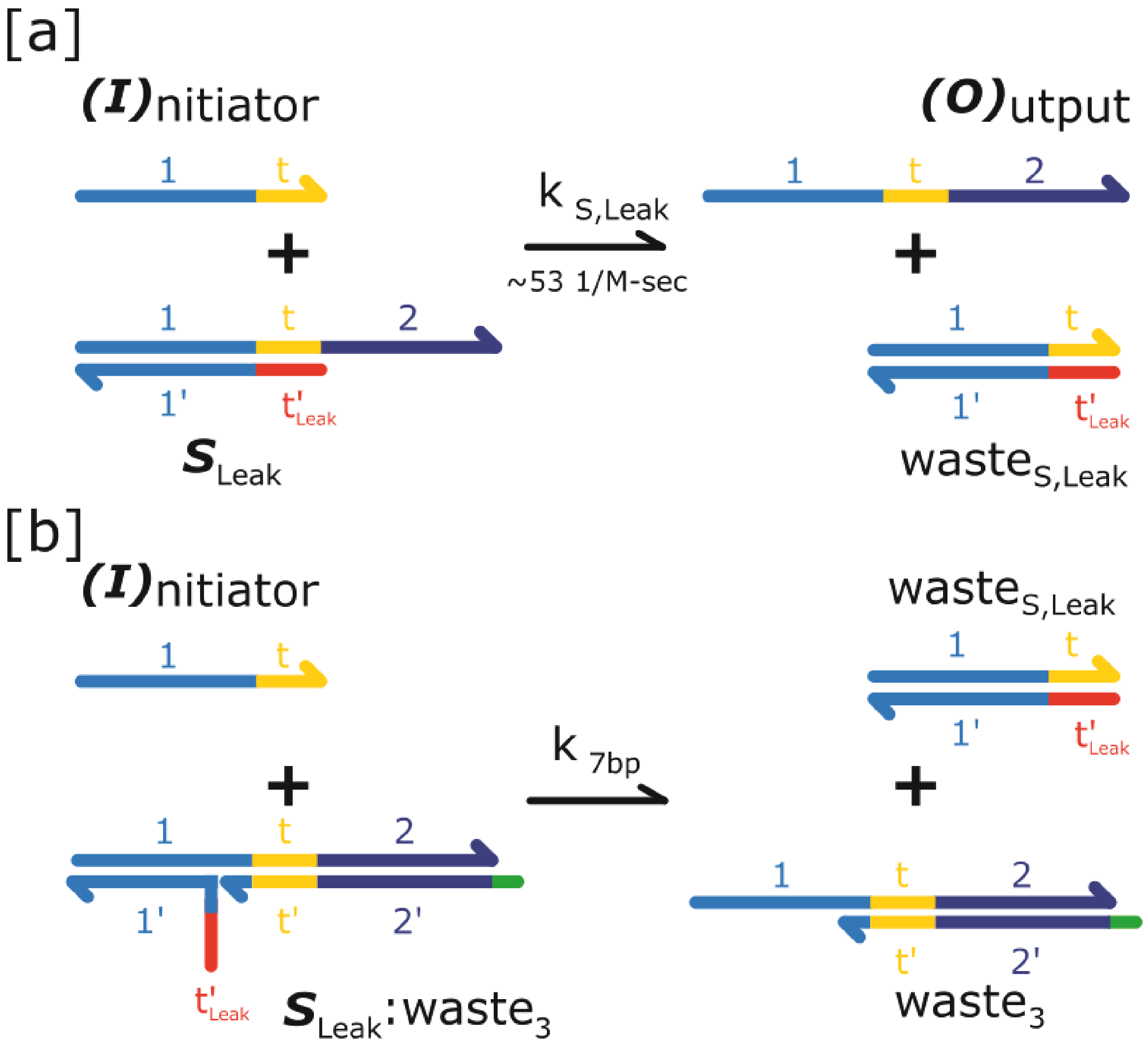


Figure S3.20: Leak reactions with Initiator. (a) Reaction between **I** and $\mathbf{S_{Leak}}$. (b) Reaction between **I** and the complex produced from $\mathbf{S_{Leak}}$+**D**.

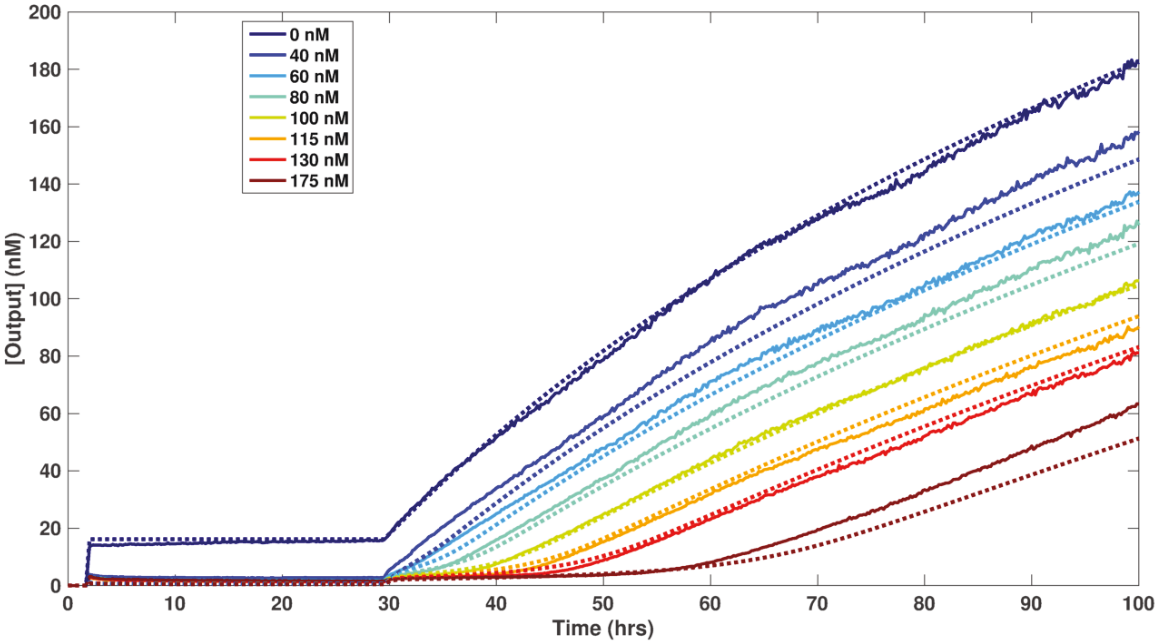


Figure S3.21: Comparison of data and the model prediction for System 1 using 1 μM **S** and **I**. Delay concentrations are listed in the legend. Reporter-only baseline was initially measured followed by Source and Delay complex addition at 1 hour. Initiator was added after 30 hours. Parameters and reactions included in the model were as described above in Figs. S3.18–20 and in Zhang and Winfree.[96]

3.5 Acknowledgements

The authors would like to thank Elisa Franco, Deepak Agrawal, Abdul Majeed Mohammed and John Zenk for insightful conversations. This work was supported by NSF-CCF-1161941, NSF-SHF-1527377, award SC0010595 from the Department of Energy for some materials and supply costs, and a grant to the Turing Centenary Project by the John Templeton Foundation.

CHAPTER 4: DESIGN AND CHARACTERIZATION OF DNA STRAND-DISPLACEMENT CIRCUITS IN SERUM-SUPPLEMENTED CELL MEDIUM

Summary

The functional stability and lifetimes of synthetic molecular circuits in biological environments are important for long-term, stable sensors or controllers of cell or tissue behavior. DNA-based molecular circuits, in particular DNA strand-displacement circuits, provide simple and effective biocompatible control mechanisms and sensors, but are vulnerable to digestion by nucleases present in living tissues and serum-supplemented cell culture. The stability of double-stranded and single-stranded DNA circuit components in serum-supplemented cell medium and the corresponding effect of nuclease-mediated degradation on circuit performance were characterized to determine the major routes of degradation and DNA strand-displacement circuit failure. Simple circuit design choices, such as the use of 5' toeholds within the DNA complexes used as reactants in the strand-displacement reactions and the termination of single-stranded components with DNA hairpin domains at the 3' termini, significantly increase the functional lifetime of the circuit components in the presence of nucleases. Simulations of multi-reaction circuits, guided by the experimentally measured operation of single-reaction circuits, enable predictive realization of multi-layer and competitive-reaction circuit behavior. Together, these results provide a basic route to increased DNA circuit stability in cell culture environments.

4.1 Introduction

Tasks across biotechnology, including drug targeting and release, *in vitro* cell culture, diagnostics, and tissue engineering could benefit from the ability to use the concentration in time and space of several different biomarkers to make decisions such as directing release or collection processes. Ideally, such decision-making processes could be performed autonomously by molecular agents or circuits as they are accomplished in living systems. Developing such methods requires augmenting molecular detection techniques with molecular circuits that can reliably operate under physiological conditions.

One potential pathway to building these systems is through the use of DNA strand-displacement reactions. Circuits that are capable of signal amplification,[109] timing control,[101] complex logic functions,[87,107] and directing the release of molecules[198] or other material control processes[144] have shown promising applications in cell culture[146] or with cells,[194,199] or for protein[200,201] and RNA[137] detection due to the robust nature of DNA sequence specificity and innately high biocompatibility. However, current DNA-based technologies have limited applications and lifetimes in blood serum and serum-supplemented cell cultures due to the presence of nucleases that degrade the DNA components.[29,138,139,202,203] Current approaches to using or protecting DNA complexes in serum[204–206] or inside cells,[136,137,207,208] are generally of limited success because they require molecular components that are difficult or expensive to synthesize, or require medium additives that can damage cells or prevent their growth. New approaches to enable strand-displacement circuits to operate in serum are therefore still needed.

In this chapter, we investigate the usage of simple and inexpensive design principles of DNA strand-displacement circuits to enable these circuits to detect concentrations, operate signal cascades, and control timing in cell culture environments over as long as 70 hours. The lifetimes of double-stranded fluorophore- and quencher-modified DNA complexes that perform these functions were measured for both 3’ and 5’ single-stranded overhangs (toeholds). Compared to 3’ toeholds, complexes with 5’ toeholds showed an increase in functional operation in the serum we tested. Additionally, single-stranded DNA species were shown to have an increased lifetime when a hairpin domain was added to the 3’ termini. DNA strand-displacement circuits designed with 5’ toeholds and 3’ hairpin domains were shown to have reaction rate constants dependent on toehold length in serum-supplemented cell media with the same rate of reaction rate constant slow down with decreasing toehold length as in standard TAE/Mg^{2+} buffer. To investigate the operation of larger, more complex DNA circuits in the presence of nucleases, we developed a model using standard enzyme reaction dynamics and DNA strand-displacement molecular reactions. The model was able to qualitatively predict the behavior of a multistage strand-displacement cascade and a timer circuit. While single-layer circuits can reliably operate for at least 12 hours, more complex circuits are affected by nuclease degradation to a greater extent, and responses were most different when the concentration of inputs is low.

4.2 Materials and Methods

4.2.1 Reagents

Fetal bovine serum (FBS, cat. no. 10437010) and penicillin/streptomycin (P/S, cat. no. 15140148) were purchased from Thermo Fisher Scientific. Dulbecco’s modified Eagle’s medium (DMEM, cat. no. D6429) and single-stranded DNA from salmon testes (STssDNA, cat. no. D7656) were purchased from Sigma-Aldrich. Chicken gizzard muscle actin protein (cat. no. AS99) was purchased from Cytoskeleton, Inc. Standard cell culture medium was prepared as DMEM supplemented with 10% FBS and 1% P/S and stored up to one month. All DNA for strand-displacement circuits was purchased from Integrated DNA Technologies, Inc., with standard desalting purification except for quencher-modified strands, which were ordered HPLC purified. Sequences of all DNA strands used here are listed in Table S4.10.

4.2.2 Preparation of DNA complexes

Source, Delay, and Reporter complexes were annealed in 1x TAE buffer supplemented with 12.5 mM Magnesium Acetate (TAE/Mg^{2+}) from 90 to 20 °C using an Eppendorf PCR at 1 °C/minute. Zero base pair toehold Source complexes were incubated 1:1 with their Initiator strand without a hairpin domain for ~16 hours prior to PAGE purification to remove some poorly formed or synthesized DNA complexes. After annealing, Source and Delay complexes were PAGE purified at 150 V for 3 or 1.5 hours using 15% or 10% polyacrylamide gels. The complexes were eluted from the excised gel band using TAE/$Mg^{2+}$and overnight incubation. Reporter complexes were not PAGE purified.

4.2.3 qPCR measurements of DNA strand-displacement reactions in cell medium

Agilent Stratagene Mx3000 or Mx3005 qPCRs were used to test the stability and lifetime of DNA strands and complexes in serum-supplemented DMEM at 37 °C. The final volume of liquid in each well was fixed at 150 μL. Of this volume, 120 – 140 μL was serum-supplemented DMEM and the remaining volume consisted of nuclease-screening components and DNA complexes and strands. Additional components were added to the media to screen the nucleases for increased protection of DNA strand-displacement circuit components. Actin protein (200 nM), 12 μM poly$(T)_{20}$ single-stranded DNA, 0.1 mg/mL STssDNA, and 10 μM of a double-stranded DNA complex with two 3' single-stranded DNA tails (mimic DNA) were mixed with the serum-supplemented DMEM immediately prior to the addition of circuit components. This mixture is referred to as nuclease-screened medium (NS medium) in the text. The fluorescence intensity of Reporter complexes alone (prior to nuclease degradation) was subtracted from all measurements (Section 4.5.4). For Release, cascade, and timer experiments, the Reporter was initially added at 200 nM, followed by the Source and/or Delay complexes. Initiator strands were added after about 1 – 4 hours post-Source addition.

4.3 Results and Discussion

4.3.1 Inhibition of nucleases using competitor molecules

In order to protect the DNA strand-displacement circuit components from potential nuclease-mediated degradation, we first investigated the use of additional biocompatible medium supplements that may act as competitive inhibitors to screen the nucleases from circuit components. In a DNA strand-displacement reaction, an invading strand binds a substrate complex and displaces an incumbent strand. The invading strand binds to a single-stranded "toehold" region of the substrate complex and undergoes branch migration with the incumbent strand of the substrate until the incumbent strand is fully displaced (Figure 4.1A). In our studies, a double-stranded DNA complex modified with a fluorophore and quencher on complementary strands, called a Reporter, was used as a basis for measuring the functional stability of the circuit components in the presence of nucleases. A 5 base pair (bp) toehold on the 5' end enabled strand-displacement to occur upon the addition of the invading strand, termed the Output, which in a more complex circuit would be the released species whose concentration would be measured by interaction with the Reporter complex (Figure 4.1A). When the Output interacts with a Reporter complex, the fluorescence increases due to the Output displacing the quencher-modified strand. Upon degradation of the Reporter complex by nucleases, the two strands of the Reporter complex may dehybridize, also causing an increase in fluorescence (Figure 4.1B). Thus, a change in the measured fluorescence of a sample containing nucleases can be due to either the reaction of the Reporter complex with an Output strand or digestion of the Reporter complex by nucleases. In DMEM with 10%

fetal bovine serum, nuclease induced degradation of the Reporter complexes causes complex dehybridization (Figure 4.1C).

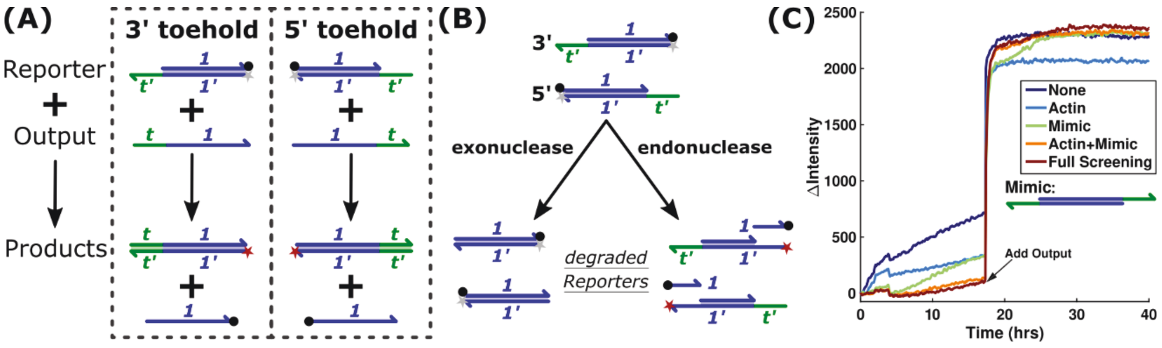


Figure 4.1: Comparison of nuclease screening supplements to protect double-stranded Reporter complexes. (A) Reporter complexes with either a 3' or a 5' toehold and their respective complementary Output strands. The top strand of each Reporter complex is modified with a quencher and the bottom strand with a FAM fluorophore. Blue domains are 13 base pairs and green toehold domains are 5 nucleotides. Domains and their complements are denoted by a shared number or letter, with the complementary domain accented with a prime (*e.g.*, 1 and 1'). See Table S4.10 for sequences. (B) Reporter complexes can be degraded by exonucleases or endonucleases. Exonucleases remove all or part of the single-stranded toehold domain, preventing Output strand hybridization at the toehold that initiates strand-displacement. Endonucleases nick the interior domain potentially causing denaturation of the complex, separation of the fluorophore-quencher, and an increase in fluorescence. (C) Output strands (200 nM) were added to 100 nM 5' toehold Reporter complexes incubated for 18 hours (arrow) in serum-supplemented medium with varying competitors for nuclease-DNA complex interactions (nuclease screening molecules). Actin was added at 200 nM, and mimic DNA, a 13 base pair double-stranded DNA complex with 5 nucleotide 3' single-stranded overhangs (inset schematic), was added at 10 μM. Full screening solution consisted of actin, mimic DNA, 12 μM Poly(T)20 single-stranded DNA, and 0.1 mg/mL single-stranded DNA from salmon testes.

Actin has previously been shown to inhibit the activity of DNases.[209–211] When actin was added to a final concentration of 200 nM to serum-supplemented DMEM, the increase in fluorescence over 18 hours from 100 nM Reporter complexes was reduced by half (Figure

4.1C). To further reduce the degradation of Reporter complexes, a second DNA complex designed to act as a competitive inhibitor for nucleases was added to the medium at 10 μM. This mimic DNA contained a double-stranded region of the same length as the region in the Reporter complex and 5 bp-long single-stranded overhangs on the 3' end of each of the sequences (see Table S4.10 for sequences). The inclusion of both this mimic and actin in the serum-supplemented DMEM further reduced the rate of degradation of the Reporter complex (Figure 4.1C). Finally, 12 μM poly(T)$_{20}$ single-stranded DNA and 0.1 mg/mL single-stranded DNA from salmon testes were added to act as competitors for nucleases that bind single-stranded DNA. Together, the inclusion of these four additives in serum-supplemented medium reduced the rate of fluorescence increase of the Reporter complex in the absence of Output (which is presumably due to nuclease degradation) by about an order of magnitude. Unless otherwise specified, all DNA strand-displacement reactions in the experiments that follow were performed in serum-supplemented DMEM containing actin, poly(T)$_{20}$, mimic DNA, and single-stranded DNA from salmon testes at the concentrations given above, which we term nuclease-screened medium or NS medium.

4.3.2 Design of DNA circuit components with increased nuclease resistance

In order to build DNA strand-displacement circuits that can reliably process signals using reaction cascades in serum-supplemented medium, we next investigated how changes to the complexes themselves could reduce the rate of nuclease digestion without eliminating the ability to use the complexes to design molecular circuits with controlled rates. The Reporter complex was again chosen as a model for the investigation of DNA circuit component design. In addition to the degradation of double-stranded domains by endonucleases (Figure 4.1C), exonucleases can digest any single-stranded region (Figure 4.1B). Since Reporter toeholds are generally placed at either the 5' or 3' termini of complexes (Figure 4.1A), the toehold is susceptible to degradation from exonucleases, leading to the loss of the toehold domain and reactivity to Output strand input over time (Figure 4.1B). For each base removed from the toehold, the reaction rate constant between the Reporter and Output species is expected to decrease by an order of magnitude, based on measurements in standard buffers.[96]

To characterize the functional operation of Reporter complexes, the maximum fluorescence of the mixtures was measured after Output strands were added to 5' or 3' toehold Reporter complexes incubated for different periods of time at 37 °C in NS medium (Figure 4.2). Reporters with 3' toeholds showed a marked decrease in the change in fluorescence upon Output addition for Reporter incubation times as short as 6 hours, indicating that the incubation of the Reporter significantly reduced its ability to interact with the Output. In contrast, Reporters with 5' toeholds showed only a 25% drop in fluorescence when the Reporter complexes were incubated in NS medium for 70 hours prior to addition of the Output strands. Further, the time to reach reaction equilibrium upon the addition of the Output strand was much slower for the Reporter with the 3' toehold as compared to the Reporter with a 5' toehold, suggesting that the kinetics of strand-displacement was affected more by incubation for the Reporter with the 3' toehold (Figure S4.1). The difference between the responses of the two types of complexes that was observed is consistent with a significantly larger concentration or activity of 3' exonucleases than 5' exonucleases present in the serum.[131,212]

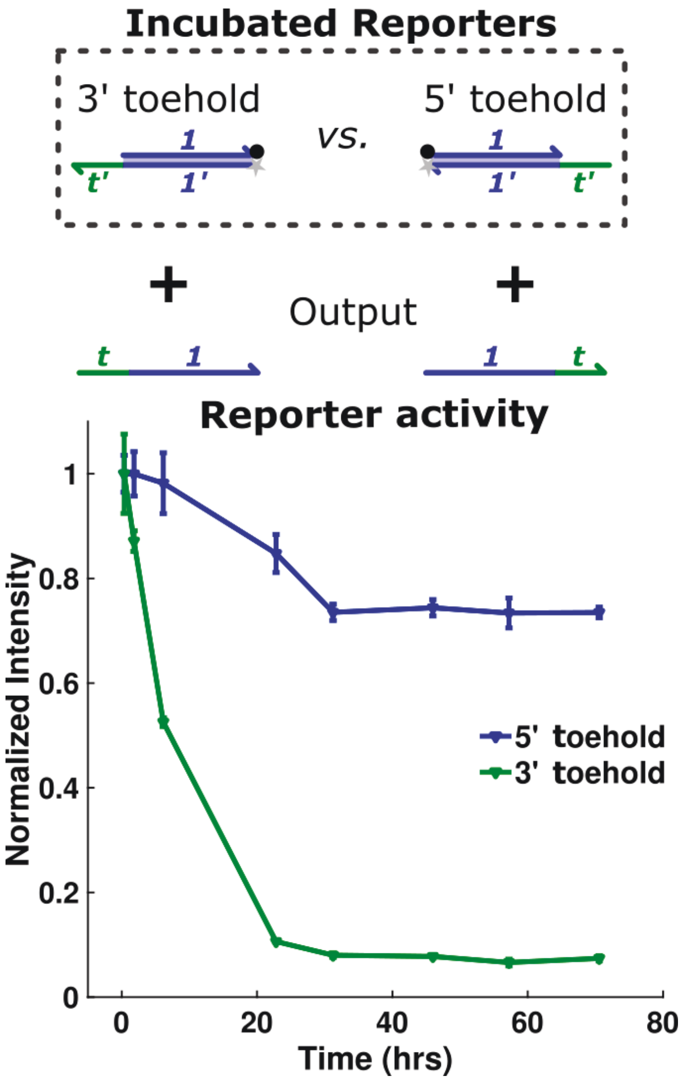


Figure 4.2: Comparison of the functional stability of Reporter complexes with either 3' or 5' toeholds in nuclease-screened serum-supplemented DMEM. Output strands (250 nM) were added to 200 nM Reporter complexes at 0.5, 2, 6, 24, 32, 48, 56, 66 and 72 hours post Reporter addition to medium at 37 °C. Reporter complexes with a 3' toehold showed a marked decrease in fluorescence intensity change when the Output was added at subsequent time points compared to the sustained reactivity of the Reporter with a 5' toehold. Both complexes show an increase in fluorescence in the absence of Output over time due to the presence of nucleases, potentially leading to the decrease in 5' Reporter reactivity that was observed (Figure S4.1). Points are relative to the intensity change upon Output addition at 0.5 hours of incubation (Section 4.5.2) and are the average value of 2 – 3 replicates. Error bars are one standard deviation of replicate results.

For strand-displacement circuits to operate reliably in serum, DNA complexes and strands that interact with or are released by the complexes (*e.g.*, Output strands) must remain intact when they are in single-stranded form rather than being degraded by nucleases. Two possible DNA modifications, phosphorothioate backbones[136,213,214] and an inverted thymine modified base[212] placed at strand termini (Figure S4.2A), have previously been shown to increase the stability of single-stranded DNA molecules against nuclease digestion. To test whether strands containing phosphorothioate bonds (PS bonds) or an inverted dT (Inv-dT) modification would remain capable of participating in strand-displacement reactions in medium for longer periods of time than unmodified strands, we designed a series of Output strands with these modifications at different sites and measured the fraction of the strands that were able to react with the Reporter to produce a change in fluorescence as a function of its initial incubation time in NS medium. Output strands containing 7 nucleotides with PS bonds at both termini and strands terminated with an Inv-dT at the 3' end showed about a 25% decrease in the concentration of strands capable of strand-displacement reactions over 7 hours of incubation in NS medium (Figure S4.2). The concentration of reactive unmodified Output strands decreased 85% in the same amount of time. Additionally, DNA strands with PS-bonds demonstrated a significant decrease in reaction rate compared to unmodified or Inv-dT modified DNA, indicating a decrease in strand invasion capabilities for DNA with PS-bonds.[136] This was further demonstrated using DNA strands fully composed of PS-bonds (Figure S4.2). Thus, the use of modified DNA backbones designed to resist nuclease degradation did not sufficiently increase the Output strand's lifetime over 7 hours of incubation in nuclease-screened medium while maintaining high reactivity with the Reporter complex.

Since the majority of exonucleases attack single-stranded DNA at their 3' termini,[131,212] DNA strands containing inert, double-stranded DNA regions on the 3' ends of the resulting complex/strand would be expected to have longer lifetimes in exonuclease-containing serum-supplemented DMEM than unmodified strands.[215,216] We thus designed a modified Output strand with a hairpin added to its 3' end (Figure 4.3), providing a terminal double-stranded region to inhibit exonuclease interactions. When we used Output strands with this additional domain at the 3' end, the stability of the Output in NS medium was greatly improved over 7 hours of incubation (Figure 4.3, Figure S4.3). However, degradation did occur at longer times, with very little of the Output being able to react with the Reporter complex after 55 hours. In contrast, 85% of the unmodified strands were no longer able to react with the Reporter after just 7 hours. This modification therefore enables single-stranded species to operate over extended time periods without chemical modification. Thus, modifying 3' termini with hairpins, the use of nuclease-screened medium, and the use of 5' toehold Reporter complexes together suggested a way by which we could design strand-displacement circuits that can operate in cell medium.

To further demonstrate the versatility of using 3' terminal hairpins to protect 3' termini from exonuclease attack, we redesigned the Reporter with 3' toeholds to include a 3' hairpin adjacent to the toehold region (Figure S4.4). We measured the Output strand activity in experiments analogous to those shown in Figure 4.2 for this modified Reporter. The addition of the 3' hairpin increased the functional lifetime of the 3' toehold Reporter to slightly less than a 5' toehold Reporter over 32 hours. Thus, DNA strand-displacement architectures that depend on free 3' termini for toehold domains and reversible reactions[87,107,109,191] could operate as reliably as those with only 5' toeholds if a hairpin motif is included at the 3' end.

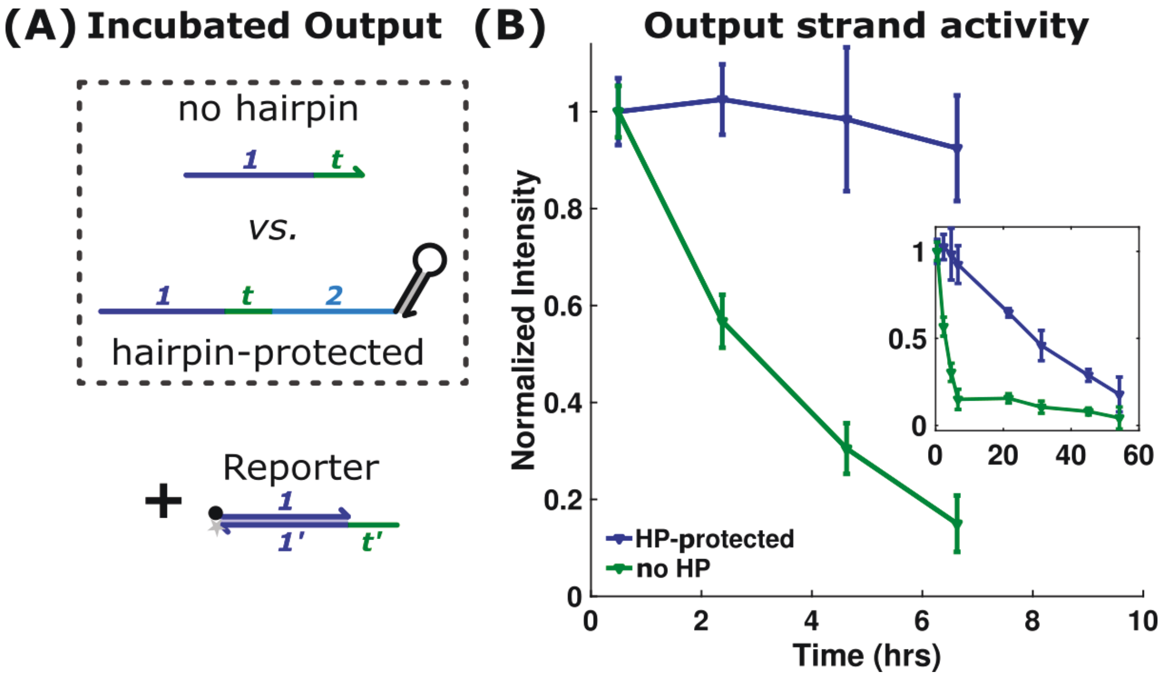


Figure 4.3: Design and characterization of a 3' exonuclease-inhibiting domain for single-stranded DNA species in nuclease-screened medium. (A) To increase stability against nucleases, Output strands were modified with a hairpin domain providing a double-stranded region at the degradation-susceptible 3' termini. The ability of the Output strand to participate in a strand-displacement process after incubation in serum was tested using the Reporter complex shown. (B) The relative concentration of 100 nM Output strands, with or without the hairpin, were measured using a 5' toehold Reporter (200 nM) that was added after different periods of Output incubation in NS medium. Output strands with the hairpin domain showed sustained stability over 7 hours of incubation. Without the hairpin domain, Output strands were largely degraded within 7 hours. Degradation is measured as the relative intensity change upon Reporter addition as compared to the intensity change after 0.5 hours of incubation (Section 4.5.2) and is the average value of 2 – 3 replicates. Error bars are the standard deviation of replicate results. *Inset:* Measurement of Normalized Intensity showing the same measurement of relative Output concentration over 55 hours of incubation.

4.3.3 Operation of simple DNA strand-displacement circuits

To investigate how well the NS medium and 3' hairpin modifications work together to protect the DNA components of strand-displacement circuits from nuclease-mediated degradation, we developed a simple DNA strand-displacement circuit, termed a Release module, that is designed to release an Output strand at a programmed rate upon the addition of an Initiator strand (Figure 4.4). In buffers such as 1x TAE/12.5 mM Mg^{2+} (*i.e.*, TAE/Mg^{2+} buffer), similar strand-displacement systems have served as the building blocks of multistage, programmable cascades for logic[87,107] and timing operations.[101] A Reporter complex captures the Output strand, and the cascade is designed so that the level of fluorescence reflects the concentration of Output at a given time. In TAE/Mg^{2+} buffer at room temperature, the rate at which Output strands are released is primarily controlled by the number of free single-stranded nucleotides in the toehold on the Source complex when the toehold ranges in size from 0 to 7 bases.[96] This ability to control reaction rates using this design principle is critical for building strand-displacement cascades with reliable responses to different inputs and timing.

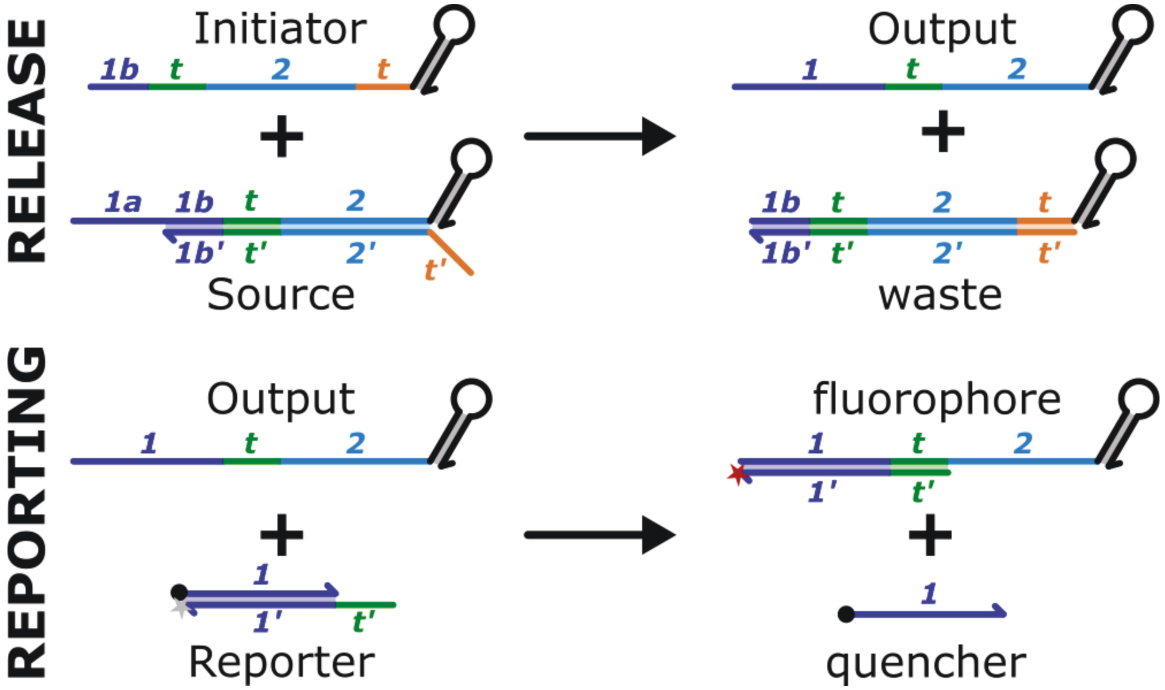


Figure 4.4: Schematics for a DNA strand-displacement reaction that releases an Output strand from a sequestered state and a reaction that produces a fluorescence output to report on the rate of release. The Output is produced when a Source and Initiator react. This reaction is initiated by a toehold ("t" domain, salmon color) of either 0, 2, or 5 bases in length presented at the 5' end. Green colored toeholds (t) are 5 base pairs, and numbered domains are 13 base pairs. The *1* domain is split into *a* and *b* parts where *b'* serves as a "clamp", or energy barrier that reduces the rate of undesired strand-displacement reactions. The black hairpin domains on the Output and Initiator strands decrease the propensity for exonuclease-mediated degradation at the 3' ends of single-stranded DNA regions without significantly affecting the rates of hybridization or strand-displacement.

To verify that the kinetics of strand-displacement reactions could also be controlled by varying the length of the toehold when the reactions occurred in NS medium at 37 °C, we measured the kinetics of Output release using Source complexes with toeholds of length 0, 2, and 5 bases. Each Source complex was combined with Initiator, and the Output strand concentration was measured as a function of time by measuring the fluorescence produced by its reaction with a 5' toehold Reporter (Section 4.5.4, Figure S4.5). The kinetics were measured using a Source concentration of 200 nM for toehold lengths of 2 and 5 bases with Initiator concentrations ranging from 0 to 300 nM. Since the kinetics were slower for 0 nucleotide toeholds, higher reactant concentrations were used. The Source and Initiator concentrations were kept equal and ranged from 0.5 to 1.5 μM. In each case, the Reporter concentration was at 200 nM. In all cases, some increase in fluorescence was observed when Reporter and Source complexes were mixed together before Initiator was added (Figure 4.5). This initial increase was higher than when the Reporter was incubated alone in NS medium (Figure S4.5). One possible reason for the increase is the degradation of Source complexes by nucleases into partially degraded, but still "active", Output strands. Since a strand-displacement reaction is expected to occur between an Output and Reporter complex so long as the Output has a toehold and a free *1* domain (Figure 4.4), some excision fragments of the Source after nuclease digestion could lead to the release of partial Output strands that are "active" because they have a contiguous toehold and *1* domain. While this reaction occurs, it is generally not the dominant pathway for Output release. For each reaction studied, the rate of Output released from strand-displacement reactions was faster than the rate of release without Initiator present, *i.e.*, through nuclease-mediated release alone.

To test whether 3' hairpins and 5' toehold design choices alone enable reliable operation of strand-displacement circuits in cell medium without the addition of the actin and DNA that functioned as a screen for nucleases, we conducted the same Release reactions using Source complexes with 5 bp or with 2 bp toeholds without nuclease-screening components (Figure S4.6). We found that, without the nuclease-screening components, the rate of digestion of Reporter and Source complexes in the absence of Initiator increased 2-fold over 40 hours of incubation. Additionally, for Source complexes with a 2bp toehold, only a minimal difference in [Disassembled Reporter] was observed for Initiator concentrations between 50 and 300 nM.

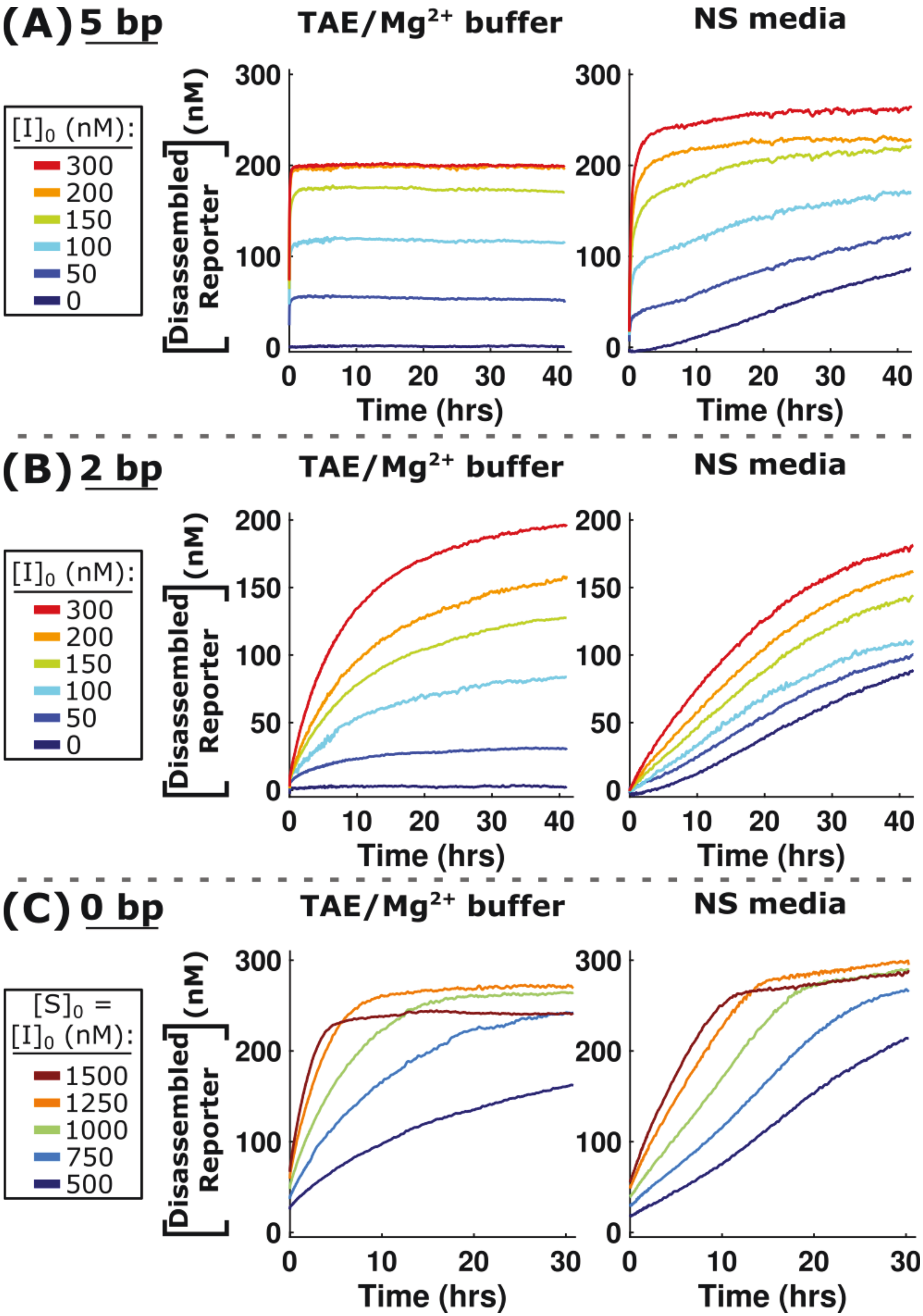

(A) 5 bp
TAE/Mg2+ buffer
NS media
[I]0 (nM):
300
200
150
100
50
0
[Disassembled Reporter] (nM)
Time (hrs)
(B) 2 bp
TAE/Mg2+ buffer
NS media
[I]0 (nM):
300
200
150
100
50
0
[Disassembled Reporter] (nM)
Time (hrs)
(C) 0 bp
TAE/Mg2+ buffer
NS media
[S]0 = [I]0 (nM):
1500
1250
1000
750
500
[Disassembled Reporter] (nM)
Time (hrs)

Figure 4.5: Measured kinetics of the toehold-mediated strand-displacement cascade shown in Figure 4.4 for Source complexes with different toehold lengths as shown. Release reactions initiated by 5 base pair (A), 2 base pair (B), and 0 base pair (C) toeholds on the Source demonstrate the range of kinetics available by varying toehold length under nuclease-screened, serum-supplemented conditions (*right*) and in TAE/Mg^{2+} buffer (*left*). (A, B) Source complexes with 5 bp or 2 bp toeholds at 200 nM were mixed with varying concentrations of Initiator as listed with each plot. The concentration of Disassembled Reporter increases over time either because the top strand of the Reporter is displaced by an Output strand or, in NS medium, because of digestion of the Reporter and Source complexes by nucleases (Figure 4.1). (C) Source complexes with 0 bp toeholds were reacted at equimolar concentrations with Initiator strands (0.5, 0.75, 1, 1.25 and 1.5 μM). In all reactions tested, the Output strand was released faster when the Initiator was present than when release was caused by the degradation of the Source complex alone (Figures S4.11 and S4.12). In all cases, the Reporter concentration was 200 nM.

4.3.4 Modeling DNA strand-displacement circuits under nuclease-degrading conditions

To design strand-displacement systems with reliable kinetics and dose-response behavior in serum, we will need to be able to predict the kinetics of both strand-displacement and degradation-mediated release of oligonucleotides. To work toward such a capacity for prediction, we developed a model of circuit kinetics (Section 4.5.5) that incorporates both standard strand-displacement reactions[87,107] and reactions for nuclease-substrate binding.[210,217] Our goal was to build the simplest possible model that is capable of predicting the kinetics we expected to observe, so our model does not necessarily completely capture the many potential degradation pathways and interactions between intermediates in the system.

The model uses mass action kinetics and makes the standard assumption that enzymes initially reversibly bind to their substrate to form an intermediate. An irreversible reaction involving the enzyme-substrate complex produces a product species and a newly freed enzyme. To reduce the number of species and potential reactions, actin and the other screening components were grouped into a general class of molecules termed "Inhibitors", which reacted with nucleases through the following composite reactions:

$$E + Inh \underset{k_{EIrev}}{\overset{k_{EIf}}{\rightleftharpoons}} E{:}Inh \xrightarrow{k_{EIDeg}} E + Waste \tag{4.1}$$

where *E* represents the combined pool of endonuclease and exonuclease enzymes, *Inh* are inhibitors, and *E:Inh* is an enzyme-inhibitor complex. While we assume in equation 4.1 that actin is degraded by nucleases for simplicity, this has not been shown to be the case thus far experimentally.

DNA complexes that were strand-displacement circuit components were assumed to interact with and be degraded by nucleases *via* a similar set of composite reactions. Each DNA complex was assumed to first be degraded by nucleases into a partial complex, with partial complexes then being degraded into waste/product molecules. This model produces the delayed degradation dynamics observed in experiments involving Reporter alone or with Source but no Initiator molecule (Figures S4.7 and S4.9 – S4.12). To model the increase in fluorescence that is caused by the separation of the FAM fluorophore and quencher molecule when the Reporter complexes are degraded, we used the reactions

$$E + Rep \underset{k_{ERrev}}{\overset{k_{ERf}}{\rightleftharpoons}} E{:}Rep \xrightarrow{k_{ERDeg}} E + Rep_{Partial} \tag{4.2}$$

$$E + Rep_{Partial} \underset{k_{ERrev}}{\overset{k_{ERf}}{\rightleftharpoons}} E{:}Rep_{Partial} \xrightarrow{k_{ERDeg}} E + FAM + Quencher \tag{4.3}$$

where *Rep* and *Rep$_{Partial}$* represent the complete and partially degraded Reporter complex, and *FAM* is a species containing a FAM fluorophore, which, because of the extent of degradation, now does not react significantly with other species. The degradation of other double-stranded DNA molecules, such as the Source complex, is modeled in the same manner as the degradation of the Reporter, except that the degradation products were either an "active" Output strand or waste DNA molecules:

$$E + Source \underset{k_{ESrev}}{\overset{k_{ESf}}{\rightleftharpoons}} E\!:\!Source \xrightarrow{k_{ESDPart}} E + Source_{Partial} \tag{4.4}$$

$$E + Source_{Partial} \underset{k_{ESrev}}{\overset{k_{ESf}}{\rightleftharpoons}} E\!:\!Source_{Partial} \xrightarrow{k_{ESDeg}} E + Waste \tag{4.5}$$

$$E + Source_{Partial} \underset{k_{ESrev}}{\overset{k_{ESf}}{\rightleftharpoons}} E\!:\!Source_{Partial} \xrightarrow{k_{ESDeg}} E + Output \tag{4.6}$$

An "active" Output strand contains at least part of the toehold domain and is capable of conducting toehold-mediated strand-displacement reactions with downstream complexes (*e.g.*, a Reporter) to completion. These equations account for the observed fluorescence intensity increase of the Source and Reporter mixtures in the absence of Initiator strands over that of Reporter complexes alone (Figure 4.5).

To model the degradation of the single-stranded DNA components of strand-displacement circuits (*e.g.*, Initiator and Output strands), we used a similar model, except that only one round of degradation was assumed to be sufficient for complete inactivation of the strand as this simpler model was sufficient to capture most of the kinetics observed in Figure 4.3:

$$E + ssDNA \underset{k_{EssDNArev}}{\overset{k_{EssDNAf}}{\rightleftharpoons}} E{:}ssDNA \xrightarrow{k_{EssDNADeg}} E + Waste \tag{4.7}$$

where *ssDNA* is a single-stranded DNA component (*e.g.*, Initiator or Output).

Finally, the DNA strand-displacement reactions were modeled as bimolecular reactions following earlier methods.[87,107] Both intact and partially degraded complexes (*e.g.*, $\mathrm{Source}_{\mathrm{Partial}}$ and $\mathrm{Reporter}_{\mathrm{Partial}}$) were assumed to react with their respective input species:

$$Output + Rep \xrightarrow{k_{Reporter}} FAM + Quencher \tag{4.8}$$

$$Output + Rep_{Partial} \xrightarrow{k_{Reporter}} FAM + Quencher \tag{4.9}$$

$$Source + Initiator \xrightarrow{k_{SInit}} Output + Waste \tag{4.10}$$

$$Source_{Partial} + Initiator \xrightarrow{k_{S_{Part}Init}} Output + Waste \tag{4.11}$$

Using the described model as a foundation, we next sought to fit the rate constants for each reaction by constraining the model using the experimental data presented in Figures 4.2, 4.3, 4.5, and Section 4.5.5. We used the experiments for the reaction involving the Reporter and Output alone (Figure 4.2, Figure S4.7) to fit reaction rate constants for equations 4.1 – 4.3, 4.8, and 4.9. We were able to find rate constants for nuclease-directed degradation and strand-displacement reactions that together closely recapitulated the reaction kinetics observed in Figure 4.3. In NS medium at 37 °C, we fitted a reaction rate constant of $8.5(\pm0.2)\mathrm{x}10^3$ 1/M-sec for the strand-displacement reaction involving the Reporter and the Output, slightly less than tenfold slower than the rate measured in TAE/Mg^{2+} at 25 °C, $5x10^4$ 1/M-sec.[96] A slower reaction rate constant in NS medium could be due to transient binding between Reporter complexes and nucleases and the lower concentration of divalent cations, but higher concentration of monovalent cations, in DMEM (1.8 mM Ca^{2+}, 0.8 mM Mg^{2+}, 5.4 mM K^+, 154.5 mM Na^+) compared to the 12.5 mM Mg^{2+} present in TAE/Mg^{2+}.[109,218,219]

To avoid fitting a large number of parameters simultaneously, we used the fitted reaction rate constants for the Reporter reaction to next fit a model involving Release reactions. We used data from Figure 4.5 to fit reaction rate constants for reactions between nucleases and Source complexes with 0, 2, and 5 base pair toeholds and found that the model was sufficient to capture most of the kinetics that were observed in experiments. In contrast to the fitted reaction rate constant between the Output and Reporter complexes, the fits to the rate constants between the Source and Output were comparable to the rate constants for the same reactions in TAE/Mg^{2+}at 25 °C (Section 4.5.5.3).[96] In general, the model and parameter fits demonstrate that interactions with nucleases in serum-supplemented medium are the primary contributors to deviations from established behavior of these circuits in TAE/Mg^{2+}, as compared to changes in temperature (*e.g.*, 25 *vs.* 37 °C) or ion concentrations.

Due to the assumptions in our model, the reaction rate parameters for nuclease-dependent reactions should be considered qualitative in nature. The unknown concentrations and relative activities of the individual nuclease subtypes and their respective interactions with each inhibitor molecule mean that the model will have to be recalibrated for different serum concentrations or formulations.

Additionally, because the measurements of changes in fluorescence due to Reporter complex digestion or interaction with the Output only indirectly capture the kinetics of nuclease-driven digestion, and the rates of these processes vary widely, the confidence intervals for some fits to reaction rates were very wide. Despite these limitations, the accuracy with which this simple model is able to capture the dynamics of most of the reactions occurring in NS medium across a large variation in strand-displacement reaction rate constants demonstrates the potential for this or similar models to be used as predictors of DNA strand-displacement circuit behavior in serum.

4.3.5 Predicting and verifying the operation of complex DNA strand-displacement circuits in medium

To understand how well the methods for reducing the effect of nucleases on DNA strand-displacement circuits scaled with complexity of DNA circuit systems, we used the model described in equations 4.1 – 4.11 to design and build two additional circuit systems, a multistage cascade and a timer, and investigate their function. These reactions were chosen because they illustrate the use of two design principles for scaling: the multistage cascade uses multiple reactions in series, while the timer circuit uses multiple reactions that occur in parallel.

In the multistage cascade, an Initiator signal interacts with a first complex to release a second signal, which then interacts with a second complex to release another signal, and so on, until within the final layer the signal reacts with a final complex to release an output. We designed and characterized the kinetics of a two-layer cascade, where an Initiator interacts with a Source at the second layer to release a strand that is the Initiator for the first layer. This Initiator then interacts with another Source to release the Output (Figure 4.6A). The kinetics of this process can then be measured by recording the change in fluorescence of a Reporter complex that the Output strand can react with.

Before experimentally characterizing the kinetics of the 2-layer cascade reactions, we used our model of strand-displacement and nuclease degradation and the fitted reaction rate constants to predict the kinetics of the cascade (Figure 4.6B). In the absence of nuclease-driven degradation, the concentration of disassembled Reporter would be expected to converge to the initial concentration of the Initiator, provided that the concentration of Source is at least as high at each step in the cascade. We observed that, in the 1-layer cascade in NS medium, the Disassembled Reporter concentration was generally higher than this value, and this trend was enhanced in the 2-layer cascade circuits. Because the concentration of disassembled Reporter is also limited by the total concentration of Reporter, the 2-layer cascade circuits had reduced sensitivity in its output fluorescence to different concentrations of input Initiator concentration compared to the single layer circuit. This is expected in the 2-layer cascade because there are two sets of Source complexes that can be degraded to trigger an output in the absence of inputs (Section 4.5.6.1). Cascades with slower kinetics of strand displacement, *i.e.*, those with a lower number of nucleotides in the toehold domain, had a narrow range of Disassembled Reporter concentrations in response to the range of tested input concentrations, presumably because there was more time for nuclease-mediated degradation to occur before the strand-displacement cascade could reach its final state. The simulations therefore predict that for circuits with slow release of Output strand over time, better schemes for preventing DNA degradation will be necessary to ensure that the dynamics of the circuit are not dominated by non-specific nuclease degradation.

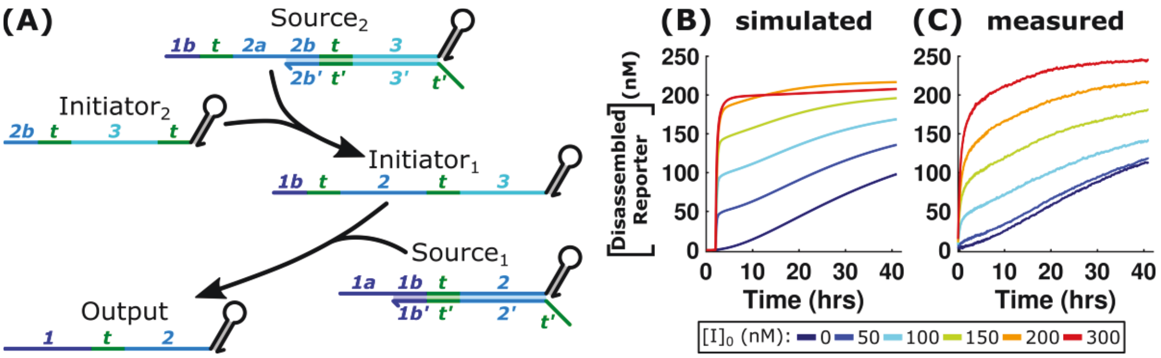


Figure 4.6: Operation of multilayer cascade circuits in cell media. (A) Schematic of the DNA complexes used in a two-layer cascade reaction. Initiator$_2$ reacts with a Source$_2$ complex to produce Initiator$_1$. Initiator$_1$ then reacts with Source$_1$ complexes to release an Output. Reporter complexes (Figure 4.4) are used to detect the rate of Output release over time. (B) Simulation of a 2-layer cascade with 5 bp toeholds using the fitted parameters as described in the text. $[I]_0$ is the concentration of Initiator$_2$. (C) Experimental operation of the 2-layer cascade circuit. Both Source complexes were at 200 nM and contained 5 bp toeholds in both the simulation and experiments. Reporter complexes are at 200 nM.

To verify the simulation's predictions, we experimentally implemented a two-layer cascade reaction (along with a reporter reaction) using 5 bp toeholds as initiation domains (Figure 4.6C). Overall, the model predicted the trend of Output release, but there were quantitative differences between the model predictions and the experimental measurements. While the model predicted that there would be significantly different rates of Reporter disassembly depending on whether 0 or 50 nM Initiator$_2$ was provided as input, there was no difference in these rates in experiments. This discrepancy suggests that the model underestimated the rate of degradation of Initiator strands, Output strands, or Source complexes.

We also developed a model for the behavior of a timer circuit that controls the delay before the release of an output begins and its rate of release (Section 4.5.6.2).[101] The circuit consists of a 0 bp Release reaction (Figure 4.4) and a delay reaction (Figure S4.17). The Release reaction slowly releases Output, which is quickly re-sequestered by Delay complexes until all Delay complexes have been reacted.[101] We used our model to predict the release kinetics of the Output for a timer circuit as a function of the concentration of the Delay complexes, which should control the length of the delay before the Output begins to be released. The simulations predicted that increasing the Delay concentration decreased the rate of Output release in the absence of Initiator, because "active" Output released from degraded Source complexes can become re-sequestered by Delay complexes. Further, we observed that changing the initial Delay concentration appeared to have only a minor effect on the timing of Output release from the circuit in the presence of Initiator. Experimental measurements of the timer circuit (Figures S4.19 and S4.20) showed that the model overpredicted the rate of Source and Delay complex degradation. Ultimately, both the experiments and model agreed that the timer circuit did not function well in serum because the Delay complex is incapable of efficiently re-sequestering Output strands after they are released from the Source complex. While, in buffers without serum, the Output is captured by Delay complexes until they are fully depleted at a reliable rate, the increased rate of release from the Source complexes, which are very abundant, and the degradation of some Delay complexes due to interactions with nucleases mean that the rate at which these parallel reactions occur is very difficult to control in serum.

4.4 Conclusion

In this chapter, we developed methods to allow DNA strand-displacement circuits to operate in cell medium supplemented with 10% FBS serum, with the goal of designing molecular systems that can detect and respond to specific concentrations and combinations of molecules by releasing different concentrations of output species. Nucleases present in the serum degraded single-stranded DNA *via* their 3' ends and double-stranded DNA complexes *via* endonuclease activity. By inhibiting nuclease activity using actin protein, and by modifying DNA complexes with hairpin extensions on the 3' ends of DNA strands, the half-life of DNA strands increased 10-fold. Using these modifications to the system, we were able to build a circuit that releases a desired Output strand with kinetics controlled by the length of the toehold domain. Fits to a model of the system showed that reaction rate constants between Initiator and Source complexes were comparable to reactions conducted in TAE/Mg^{2+} at 25 °C despite the additional DNA-binding enzymes in serum that could reduce the effective reaction rate constant and the difference in temperature. By including interactions between nucleases and DNA circuit components, the model was able to predict the dynamic behavior of Output strand release in one- and two-layer circuits. The use of models lays the groundwork for the design of more complex circuits going forward.

While we demonstrate the predictability of the DNA strand-displacement circuits using the presented model, the behavior of the circuits is still far from ideal for quantitative applications due to the strong interactions with nucleases that causes a significant loss of circuit material over time. One potential solution to enable applicability in *in vitro* cell culture experiments is to use serum-free medium, or KnockOut Serum Replacement or heat-inactivated serum as a medium supplement. However, the feasibility of both those methods largely depends on the cell type of interest and the deviation of cell behavior due to variation from the more commonly used FBS-supplemented conditions. Nucleases are expected to be present in both *in vitro* and *in vivo* settings due to nucleases secreted from dying cells or those present in blood serum, supporting greater relevancy of testing DNA strand-displacement circuits under nuclease-present conditions. It is important to note, however, that the strategy developed here of using a nuclease-screened medium is not applicable for *in vivo* devices. Implementing strand-displacement processes in whole organisms or within cells requires mechanisms for robust protection of the components of the cascade. Potential routes to such protection included the coupling to nanoparticles[212] or larger DNA nanostructures[29,137–139,203,211] that can act as steric inhibitors to nuclease interactions prior to interaction with tissues and the desired target. These limitations suggest that further methods of reducing the rate of degradation of strand-displacement circuit components due to interactions with active nucleases will still be of interest.

One limitation of this study is that we did not characterize the behavior of circuits in the presence of the CO_2 that would be required for *in vitro* cell culture. While it will be useful to verify that the circuits still function under these conditions in future studies, CO_2 is not expected have a significant impact on the operability of the DNA circuits in the absence of cells because the pH does not significantly change over the timespan of the experiments. However, circuit behavior is expected to be influenced by the presence of cells in future experiments since the cells will take up some portion of the circuit components. It will therefore be of interest to consider how the circuits we have designed behave in the presence of living cells. The ability to operate DNA strand-displacement circuits in a quantitative, predictable, and reliable manner in cell medium enables the further extension of DNA-based circuits to biological regimes yet to be studied.

4.5 Supplemental Information

4.5.1 Measuring the degradation of double-stranded DNA circuit components

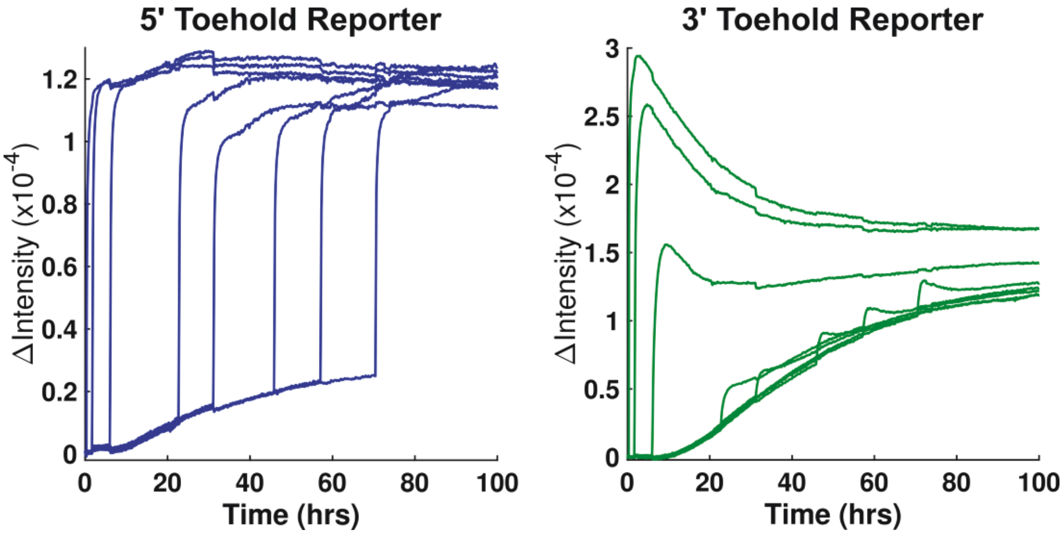


Figure S4.1: Reporter complexes (200 nM), with either a 5' (*left*) or a 3' (*right*) toehold domain, were incubated in nuclease-screened media. Output strand, with a final concentration of 250 nM, was added after 0.4, 1.8, 6.1, 22.8, 31.2, 46, 57.2, and 70.6 hours. The max intensity changes shown here are reported in Figure 4.2. Curves are the average of two – three repeats of the same experiment performed in separate qPCR wells using the same batches of materials.

4.5.2 Calculating normalized fluorescence intensity change

In order to compare the different changes in fluorescence intensity observed by reactions with 5’ and 3’ toehold Reporter complexes, and to compare the changes in fluorescence intensity when the different types of Output strands (*e.g.*, 5’ toehold, 3’ toehold, backbone-modified) are added, the change in fluorescence intensity observed when Output and Reporter were mixed together was normalized as a function of time.

This normalized intensity, used in Figures 4.2, 4.3 and Figure S4.2, was calculated using the equations:

$$Normalized\ Intensity\ = \frac{\Delta I(i)}{\Delta I(i = 0.5\ hrs)}$$

$$\Delta I(i) = Int(\ Inv.Region) - Int(Baseline)$$

where *i* represents the time point of Reporter or Output addition to incubating Output or Reporter, respectively (*e.g.*, the times listed in Figures S4.1 and S4.2). The *Baseline* is the fluorescence intensity produced from Reporter-containing solution in the absence of invading Output strand. For experiments where Reporter is incubated in NS medium, the *Baseline* is the average intensity of the measurements 20 minutes prior to adding Output strand up to the time point of Output addition. For experiments where Output was added prior to the Reporter, the *Baseline* is the average intensity of NS medium containing only Reporter over a 30 minute span immediately after Reporter was added to cell medium. The *Inv. Region* is the average intensity over the time 30 – 50 minutes after Reporter and Output are mixed together at each time point. This range was chosen as it gives enough time for Reporter and Output to react, but is before nucleases can significantly digest Reporter complexes into a disassembled state.

4.5.3 Measuring the degradation of single-stranded DNA circuit components

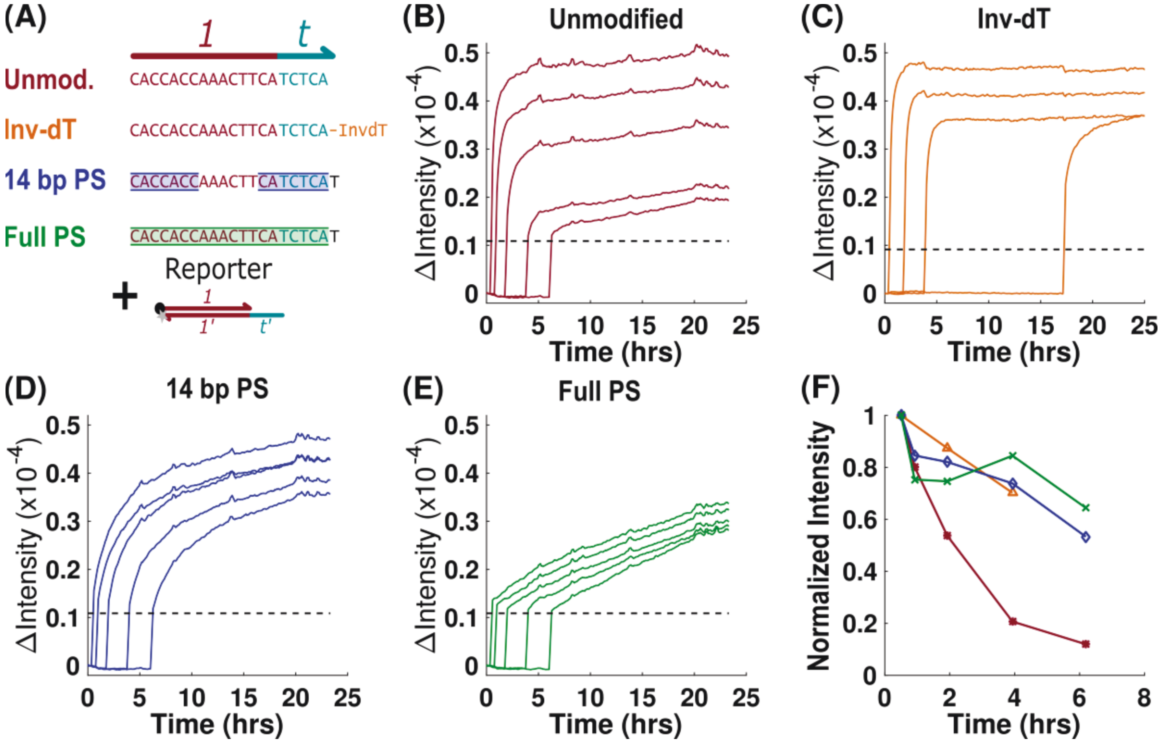


Figure S4.2: Interaction of Reporter and Output strands with and without backbone or base modifications after different periods of Output strand incubation in nuclease-screened serum. (A) Schematic of unmodified, inverted thymine (Inv-dT)-modified, or phosphothiorate-modified Output strands. An Output strand was modified with an inverted dT base at the 3' end (orange text). Phosphothiorate bonds were added along the backbone of 14 nucleotides (7 each side, blue shaded regions) or along the full length of the strand (20 nucleotides, green shaded regions). A thymine base (unmodified backbone) was added to the 3' end of each phosphothiorate-modified strand due to synthesis restrictions from IDT. (B) Unmodified Output strands were incubated at 200 nM for 0.5, 1, 2, 4, or 6 hours, after which 100 nM Reporter complex was added. The major decrease in fluorescence change in samples that had been incubated for a long time as compared to shorter times implies Output strands were mostly degraded after 6 hours of incubation. (C) The experiment in (B) was repeated with the inverted dT-modified Output with the same durations of incubation, except a 17 hour timepoint was exchanged for the 6 hour timepoint. By comparison, there is just a 25% loss of response to the Reporter after 18 hours of Output incubation in serum. (D) The experiment in (B) was repeated with the Output with 14 phosphorothioate-modified nucleotides. The response to the Reporter was 40% smaller after a 6 hour incubation in serum, suggesting that about 40% of the strands were significantly degraded by that time. (E) The experiment in (B) was repeated with the Output with only phosphorothioate-modified nucleotides. Though the absolute magnitude of the response to the Reporter complex was largely unchanged after incubation in serum, the kinetics of the reaction appeared to be significantly slower than unmodified DNA, suggesting that these modifications would not allow effective strand-displacement in serum. (F) Comparison of the relative amount of degradation observed for each modified or unmodified Output species as a function of incubation time in serum, as measured by the decrease in response to the Reporter complex. The inverted dT modification and phosphorothioate backbone modifications both significantly reduce the rate of degradation, but the phosphorothioate modifications reduce the rate of strand-displacement kinetics as well. The calculation of normalized intensity is described in Section 4.5.2.

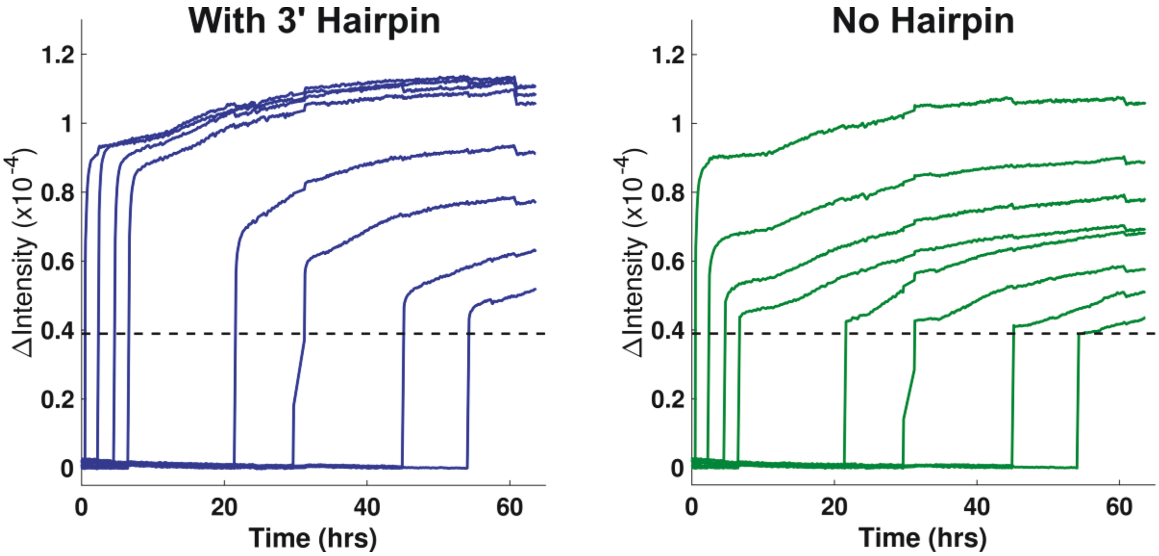


Figure S4.3: Reporter complexes at 200 nM were added at various times to 100 nM Output strands, either with (*left*) or without (*right*) a 3' hairpin domain (Figure 4.3A), incubated in nuclease-screened medium. The intensity changes seen here are normalized and reported in Figure 4.3. The dashed black line indicates the fluorescence intensity of 200 nM Reporter alone. Curves are the average of three repeats of the experiment.

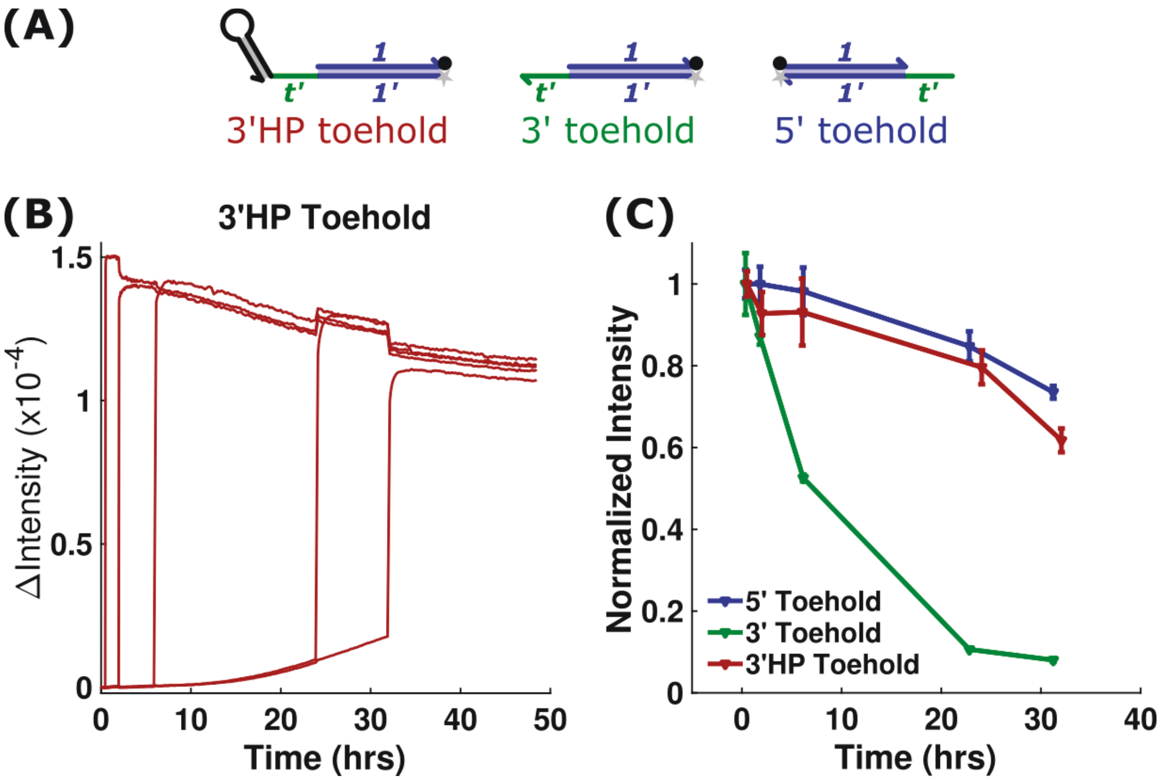


Figure S4.4: The functional stability of Reporter complexes with a 3’ toehold and hairpin in nuclease-screened serum-supplemented DMEM. (A) Diagram of the 3’ toehold Reporter modified with a 3’ hairpin and unmodified 3’ and 5’ toehold Reporters. (B) The change in fluorescence over time of 200 nM of the modified Reporter complexes in nuclease-screened serum-supplemented DMEM. Output strands (250 nM) were added to the Reporter complexes at 0.5, 2, 6, 24, and 32 hours of incubation. Each curve is the average of either two or three repeats of the same experiment performed in separate qPCR wells using the same batches of materials. (C) The change in fluorescence upon the addition of Output plotted as a function of the Reporter’s incubation time. Data for the 5’ and 3’ toehold Reporters in analogous experiments is copied from Figure 4.2 for comparison purposes.

4.5.4 Calibration of [Disassembled Reporter]

The fluorescence intensity change due to the dehybridization of the Reporter complex with 5' toeholds caused either by invasion by the Output strand or by nuclease digestion was converted into the concentration of Disassembled Reporter using a calibration curve that related known changes in hybridized Reporter complex to measured changes in fluorescence (Figure S4.5). To build this calibration curve, Output was added to a final concentration of 50, 100, 150, 200, or 250 nM to 200 nM Reporter complexes incubated in nuclease-screened medium for 30 minutes. The change in fluorescence intensity for each sample was then calculated by subtracting the average fluorescence intensity of the sample over a 10 minute period immediately prior to Output addition from the average fluorescence intensity over a 2 hour period well after the reaction between the Output and the Reporter complex had reached completion (Figure S4.5), *i.e.*:

$$\Delta Fluorescence = Avg.Intensity(t = 1.7 - 3.7\ hrs) - Avg.Intensity(t = 20 - 30\ min)$$

Based on the observed lifetime of Reporter complexes in NS medium (Figures 4.1 and 4.2), we assumed that the change in fluorescence intensity observed in these experiments was due primarily to reactions between Output strands with Reporter molecules, rather than digestion of the Reporter by nucleases. We plotted the change of fluorescence as a function of the concentration of Output that was added and used the slope of a linear fit to this plot to determine the relationship between the amount of Output added and the change in fluorescence to determine the concentration of Disassembled Reporter in our reactions using the change in fluorescence intensity.

In an ideal strand-displacement reaction, the fluorescence would reach a maximum once the concentration of Output that was added equaled the concentration of the Reporter. However, we observed that the fluorescence intensity when 250 nM of Output was mixed with 200 nM Reporter complexes was larger than the fluorescence intensity when 200 nM of Output was mixed with 200 nM of Reporter in NS medium. One potential reason for this lack of saturation could be that nucleases may bind to some reactants and make them inaccessible and thus unable to react quickly, even if they are not degraded. If more Output undergoes this process than Reporter, then we would observe that more Output than Reporter would be needed to achieve a maximum signal. Alternatively, when the concentration of Output is close to or exceeds the concentration of Reporter, virtually all Reporter must react in order to reach the expected equilibrium state. The approach to the equilibrium state may thus be limited by partially digested Reporters with shorter toeholds.

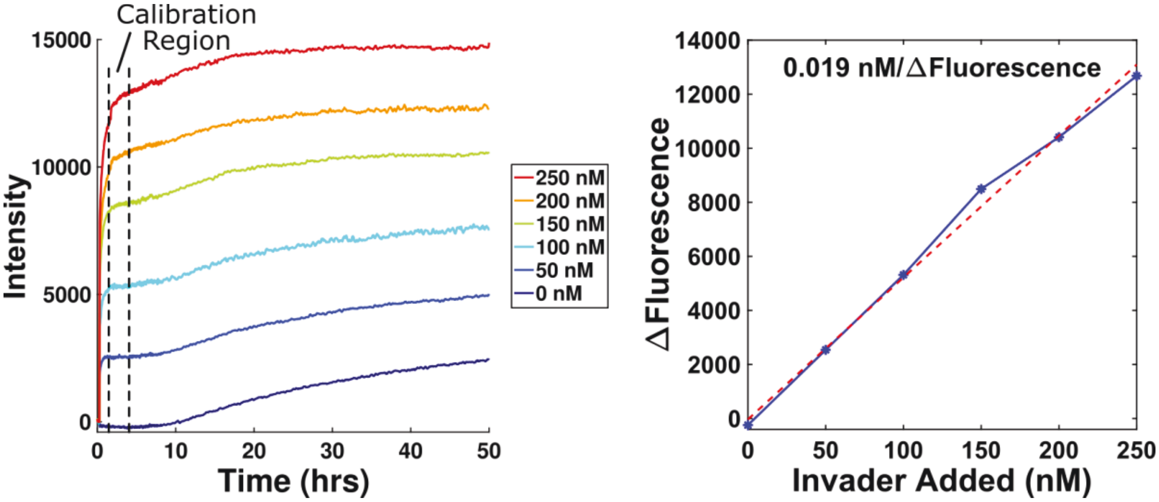


Figure S4.5: Converting the change in fluorescence intensity into [Disassembled Reporter]. The change in fluorescence in the *right* plot was calculated by subtracting the average intensity of the solution prior to Output invader addition from the average intensity of each curve in the Calibration region (*left*). The Reporter concentration is 200 nM.

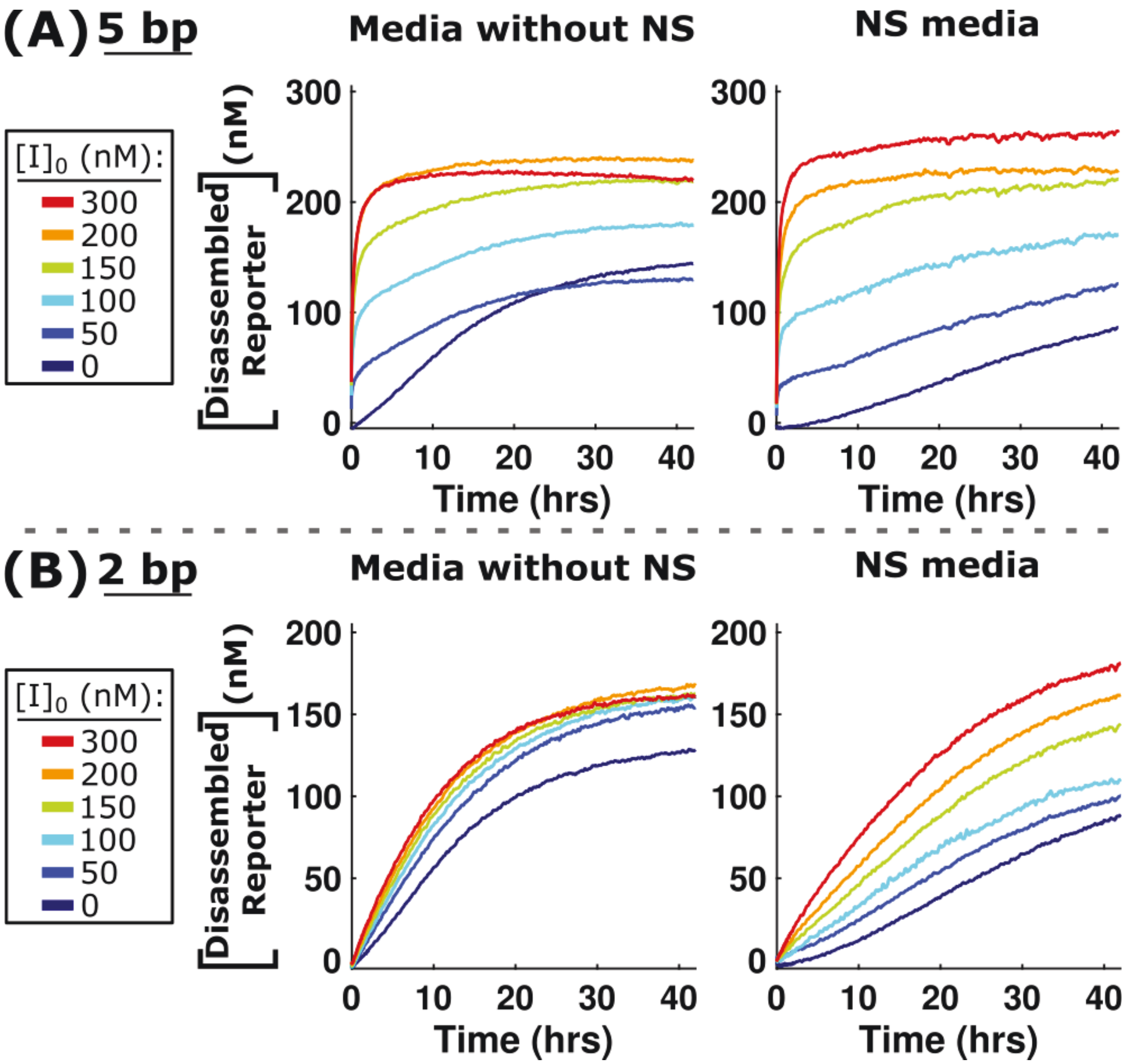


Figure S4.6: Release reactions conducted in serum-supplemented medium without nuclease-screening components (left) compared with the same experiments performed in NS medium (right). The Reporter and Source concentrations are 200 nM. The data shown for NS medium is the same as in Figure 4.5.

4.5.5 Modeling DNA strand-displacement in serum

In order to understand and predict the dynamics of the strand-displacement reactions in the presence of interfering and digesting enzymes (*e.g.*, nucleases), we generated a model that incorporates both the strand-displacement reactions and the reactions between the nucleases and the added DNA circuit components. The model also includes reactions between the nucleases and the inhibitor and screening molecules in the nuclease-screened medium that we developed.

In general, the conversion of a substrate to product using an enzyme catalyst was modeled using the standard enzyme reaction model:

$$Substrate + Enzyme \rightleftharpoons Enzyme\!:\!Substrate \rightarrow Enzyme + Product \qquad (S4.1)$$

where the *Substrate* is either a nuclease inhibitor (*i.e.*, actin or competitor DNA) or DNA circuit component, and *Enzyme:Substrate* indicates an enzyme-bound substrate intermediate complex. Here, we layout the reactions and data that was fit to obtain the figures in the main text and the estimated reaction rate constants and component concentrations.

All fitting and simulations were conducted using MATLAB's built-in functions *lsqnonlin*, *nlpredci*, and *nlparci*: standard tools for non-linear regression. Confidence intervals on fitted parameters were calculated using *nlparci,* and *nlpredci* was used to calculate 95% confidence intervals on the model predictions. For all reactions involving enzymes, the initial guess of the reaction rate constant was 0.5 1/M-sec for bimolecular reactions or 0.5 1/sec for unimolecular reactions.

As seen in Tables S4.1 – S4.6, the initial concentration of some reaction components used in fitting the reaction rate constants, Inhibitor concentrations, and nuclease concentrations were adjusted in order to obtain more sensible parameter fits. The changes that were made presumably reflected experimental variation in pipetting. Additionally, the fluorescence of the reaction mixtures usually increased beyond the expected limit that should be observed from 200 nM Output reacting with 200 nM of Reporter, and there was no way for the model to account for this discrepancy through a choice of reaction rates (Section 4.5.4). We therefore accounted for this phenomenon in the simulations by adjusting the effective Reporter concentration of each reaction mixture. Due to this adjustment and the fact that the effective concentrations of Inhibitors and Enzymes are unknown, the reaction rate constants for all reactions involving enzymes should be viewed as predictive estimates relative to the assumed concentrations in each experiment below.

4.5.5.1 Reporting reaction

We began to develop the model using a simple system containing only the Output and the Reporter complex. We used this process to fit the reactions between Output strands and the Reporter complex, the interactions between each component, and nucleases and between the nuclease inhibitors and nucleases in the nuclease-screened medium. We used a system in which the Output binds to the Reporter *via* a 5 base-pair toehold that initiates the strand-displacement process (Figure 4.1A). Since there is only a 0 bp toehold on the other side of the complex, this reaction is assumed to be irreversible due to the ~10^5-fold higher forward reaction rate constant.

$$Output + Reporter \xrightarrow{k_{Reporter}} Fluorophore + Quencher \qquad (S4.2)$$

To account for the reactions between the Reporter and the nucleases present in serum-supplemented medium, we built a model in which nucleases bind to the Reporter complex and produce a degraded product that is a semi-stable, partial Reporter complex. Partial Reporter complexes can be further digested to a fully disassembled state. We developed this model based on observations of changes in fluorescence of Reporter complexes when the Reporter is added to serum-supplemented medium. An increase in fluorescence, which in the absence of invading Output would be driven by irreversible separation of the FAM and the quencher on the complementary strand, was not immediate and constant. Instead, fluorescence increase occurred with sigmoidal-like dynamics. The delayed onset of fluorescence increase could be caused by a need for multiple rounds of degradation of the Reporter complex to occur before the FAM molecules and the quencher on its complement are no longer co-localized by hybridization. This is sensible because degradation would have to occur close to the FAM-modified termini for the remaining DNA strand to melt off the complex, resulting in a free and active FAM molecule. This multi-step process of degradation is modeled using the following reactions:

$$Reporter + Enzyme \xrightarrow{k_{ERf}} Reporter{:}\,Enzyme \qquad (S4.3)$$

$$Reporter{:}\,Enzyme \xrightarrow{k_{ERrev}} Reporter + Enzyme \qquad (S4.4)$$

$$Reporter{:}\,Enzyme \xrightarrow{k_{ERDeg}} Enzyme + Rep_{Partial} \qquad (S4.5)$$

$$Rep_{Partial} + Enzyme \xrightarrow{k_{ERf}} Rep_{Partial}{:}\,Enzyme \qquad (S4.6)$$

$$Rep_{Partial}{:}\,Enzyme \xrightarrow{k_{ERrev}} Rep_{Partial} + Enzyme \qquad (S4.7)$$

$$Rep_{Partial}{:}\,Enzyme \xrightarrow{k_{ERDeg}} Enzyme + Fluorophore + Quencher \qquad (S4.8)$$

$$Output + Rep_{Partial} \xrightarrow{k_{Reporter}} Fluorophore + Quencher \qquad (S4.9)$$

where *Rep*$_{Partial}$ is a partially degraded Reporter complex with the fluorophore in a quenched state. Because the reactions in equations S4.6 – S4.8 are based upon the same general reaction progression as in equations S4.3 – S4.5, and to reduce the total number of fitted reaction rate constants, the reactions in equations S4.6 – S4.8 were assumed to have the same reaction rate constant as their corresponding reaction in equations S4.3 – S4.5 (*e.g.*, equation S4.3 and S4.6 have the same reaction rate constant). Equation S4.9 represents the capability of Output strands to bind to partially degraded Reporter complexes and was assumed to have the same reaction rate constant as non-degraded Reporter (equation S4.2). We neglected interactions between Output strands and nucleases in cases where Output strands are added to incubated Reporters because the timescale of single-stranded degradation in nuclease-screened medium appeared significantly slower than the timescale for the Reporter-Output reaction to reach completion (*i.e.*, all Output strands bind and react with Reporter complexes). This reaction runs to completion in practice in 2 – 3 hours (Figure S4.5), whereas Outputs with 3' hairpins are resistant to degradation over 6 hours (Figure 4.3B, Figure S4.3).

To account for the effects of nuclease inhibitors in the nuclease-screened medium, we included reactions in which the enzymes (*i.e.*, nucleases) interacted with the inhibitors:

$$Inhibitor + Enzyme \xrightarrow{k_{EIf}} Enzyme{:}Inhibitor \qquad (S4.10)$$

$$Enzyme{:}Inhibitor \xrightarrow{k_{EIrev}} Enzyme + Inhibitor \qquad (S4.11)$$

$$Enzyme{:}Inhibitor \xrightarrow{k_{EIDeg}} Enzyme + Waste \qquad (S4.12)$$

Here, all inhibitors were lumped into a class of composite reactions for simplicity because the exact, combined effects of the actin protein and the inert DNA strands on nuclease activity in nuclease-screened serum-supplemented medium are unknown. While reaction rate constants for some of these reactions involving specific nucleases (*e.g.*, DNaseI) have previously been measured for models implementing the Michaelis-Menten approximation,[131,210,217] it was assumed that the varying and unknown types, and concentrations, of nuclease subtypes within fetal bovine serum, and the complexity of inhibitor types used, precluded the use of such an approximation and corresponding parameters.

The reaction rate constants for the above reactions (equations S4.2 – S4.12) and the concentrations of the Inhibitors and Enzymes were fit using the measured kinetics of the Reporter and Output strands in nuclease-screened medium (Figure S4.5, fits shown in Figure S4.7). The fitted rate constants and concentrations are listed in Table S4.1. These reaction rate constants and Enzyme/Inhibitor concentrations were assumed in fitting further experiments in order to expand the model.

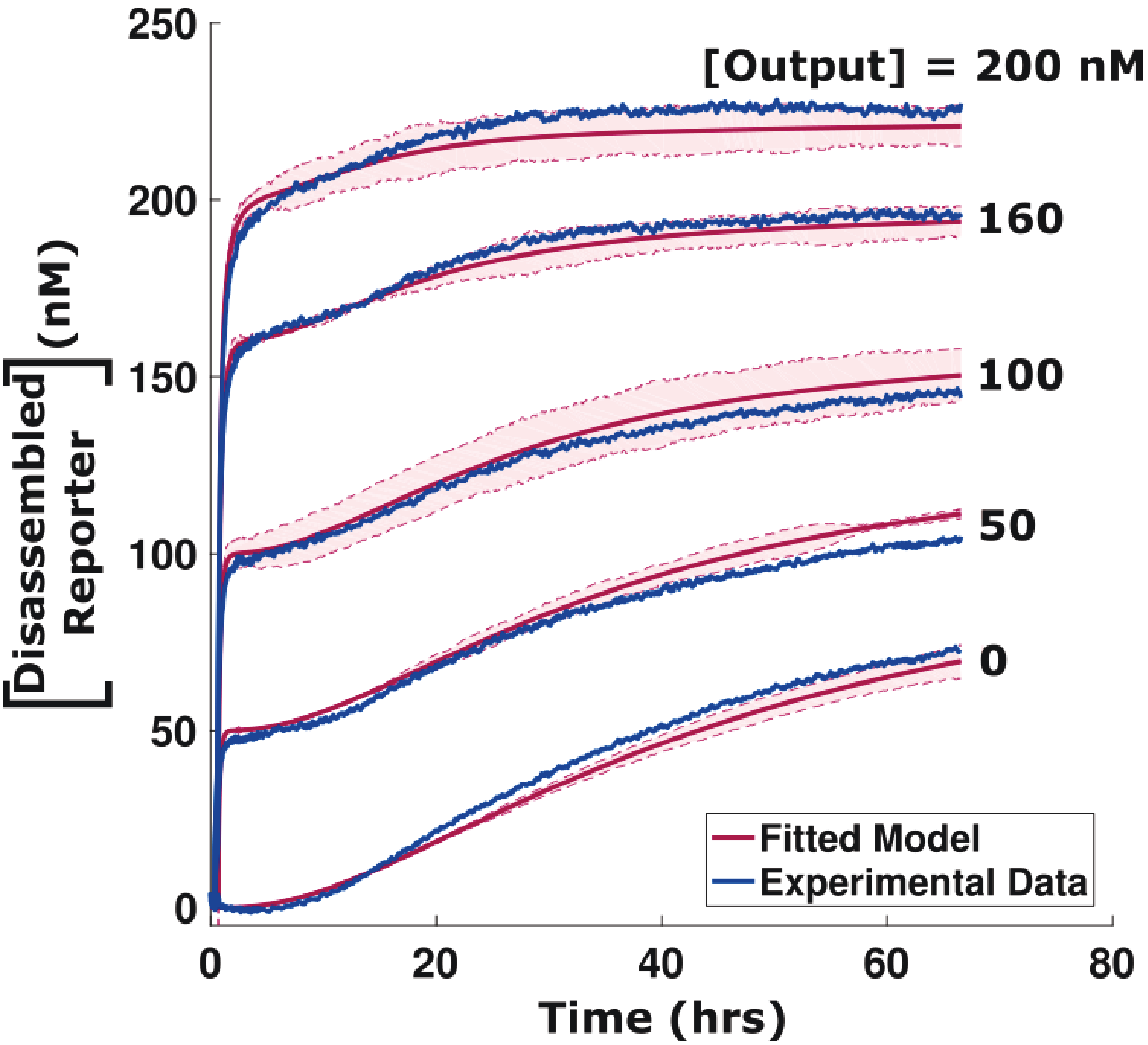


Figure S4.7: Experimental characterization of the change in fluorescence when Reporter (5 nucleotides, 5' toehold) complexes and Output strands are combined in nuclease-screened medium (same data as in Figure S4.5) and fitted using the model described in Section 4.5.5.1 with parameters as given in Table S4.1. Output concentrations listed are the simulated concentrations. Shaded red regions show the range of values predicted with 95% confidence.

Table S4.1: List of parameters used for fitting the model for the irreversible reporter in nuclease-screened medium and the calculated fitted parameters. Parameters are listed as the fitted value +/- their 95% confidence interval in parentheses. Some constants have large error values. This large range of valid parameters may be caused by the fact that varying these rates does not vary the outcome significantly. For example, the model appears to fit the data for a wide range of values of k_{ERrev}, so long as it is smaller than the rate of degradation, *i.e.*, most substrate that binds to a nuclease is degraded. We observed that the model did not work well without including these parameters and reactions.

Reaction Component	Expt. Conc. (nM)	Modeled Conc. (nM)	Rate Constant	Fitted Parameter
Reporter	200	250	$k_{Reporter}$ (1/M-sec)	$8.5(0.2)\text{x}10^{3}$
Output	0, 50, 100, 150, 200	0, 50, 100, 160, 200	k_{ERf} (1/M-sec)	$7(3)\text{x}10^{3}$
Enzyme		31.3	k_{ERrev} (1/sec)	$4(2\text{x}10^{9})\text{x}10^{-14}$
Inhibitor		620	k_{ERDeg} (1/sec)	$8.6(0.8)\text{x}10^{-5}$
			k_{EIf} (1/M-sec)	$3(1.5)\text{x}10^{2}$
			k_{EIrev} (1/sec)	$5(1000)\text{x}10^{-7}$
			k_{EIDeg} (1/sec)	$2(1\text{x}10^{7})\text{x}10^{-11}$

4.5.5.2 Single-stranded DNA degradation

In addition to DNA complexes, the single-stranded DNA components are also susceptible to degradation by nucleases. This effect is observed in Figure S4.2 and was one of the primary influencers in the choice of strand design, especially the hairpin domain added to the 3' termini of all circuit strands expected to be in a single-stranded form (*i.e.*, Output and Initiator strands). To attempt to quantify the rate and degree of degradation of Output strands with 3' hairpin domains due to nucleases for these DNA strand-displacement circuits, we used the data shown in Figure S4.3, the reactions in equations S4.2 – S4.12 and the parameters and concentrations in Table S4.1 to fit the rates for the following reactions:

$$Output + Enzyme \xrightarrow{k_{EssDNAf}} Output{:}\,Enzyme \qquad (S4.13)$$

$$Output{:}\,Enzyme \xrightarrow{k_{EssDNArev}} Output + Enzyme \qquad (S4.14)$$

$$Output{:}\,Enzyme \xrightarrow{k_{EssDNADeg}} Enzyme + Waste \qquad (S4.15)$$

The parameters that were fit are show in Table S4.2 and the use of these fits to model the process is shown in Figure S4.8. Although the confidence intervals for the fit parameters are quite small, the overall fit to the experimental data shows significant deviation between experiment and model for long Output strand incubation times, especially at 45 and 55 hours. Re-fitting the Reporter-Enzyme reactions concurrently with the Output-Enzyme reactions did not provide an improved fit to the data at long times. The lack of agreement between the model and the data here indicate that there are further interactions occurring in the experiment that we have not accounted for with the model.

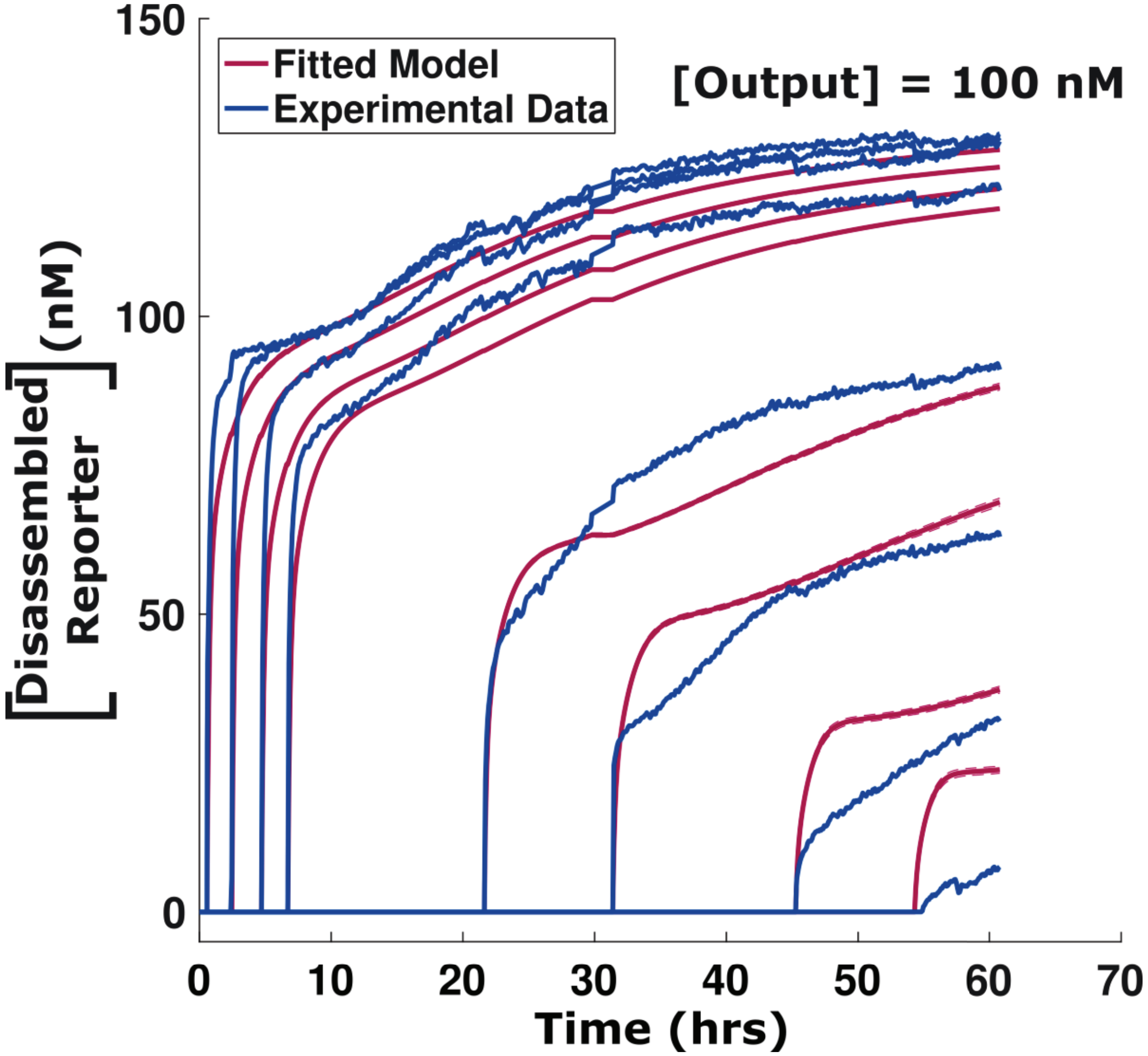


Figure S4.8: The predictions of our model for how Output strands respond to the Reporter complex after the Output is incubated in nuclease-screened medium for different times. Reporter complexes at 200 nM were added at various times to 100 nM Output strands that contained a 3' hairpin domain. The fitted model has parameters determined as described in Sections 4.5.5.1 – 4.5.5.2. Confidence intervals (95%) for the fit are too small to be observed. Fitted reaction rate constants and modeled concentrations are listed in Table S4.2. One limitation of our model is that there is only one rate at which the Output and Reporter can react, even if the Reporter or Output are partially degraded. The model is therefore unable to fit cases where the concentration of disassembled Reporter slowly rises due to a slow reaction rate between partial Output and Reporters.

Table S4.2: List of fit values for the parameters for the model for the degradation of single-stranded DNA components in serum-supplemented medium as described in Section 4.5.5.2 using the data shown in Figure S4.8. Parameters are listed as the fitted value +/- their 95% confidence interval in parentheses.

Reaction Component	**Expt. Conc. (nM)**	**Modeled Conc. (nM)**	**Rate Constant**	**Fitted Parameter**
Reporter	200	200	$k_{EssDNAf}$ (1/M-sec)	$1.69(0.07)\text{x}10^{7}$
Output	100	100	$k_{EssDNArev}$ (1/sec)	$9.0(0.8)\text{x}10^{-3}$
Enzyme		31.3	$k_{EssDNADeg}$ (1/sec)	$1.64(0.02)\text{x}10^{-5}$
Inhibitor		620		

4.5.5.3 Release reactions

We used the model developed in Sections 4.5.5.1 – 4.5.5.2 as a basis for predicting and fitting more complex strand-displacement systems for use in serum and the nuclease-screened medium we developed. In standard buffers, such as TAE supplemented with Mg^{2+}, the rate constants for toehold mediated strand-displacement processes, such as the Release reaction (Figure 4.4) are determined primarily by the length of the toehold domain of the Source complex.[96] In TAE/Mg^{2+}, the reaction rate constant increases by a factor of 10 for each base added to the toehold for toehold lengths less than 7 bases. To determine whether this rule-of-thumb applies to DNA strand-displacement reactions conducted in serum-supplemented medium and at 37 °C, we fit rate constants for strand-displacement processes between a Source complex and Initiator strand (Figure 4.4) involving toeholds of 0, 2, and 5 base pairs on the Source complex. The release of Output strand was monitored using the fluorescence of an irreversible reporter (Figure 4.4). The degradation of the Source complex was modeled in a fashion similar to the degradation of the Reporter complex – degradation was assumed to occur in multiple stages. For simplicity, we assumed that the reaction rate constants governing the nuclease-driven degradation of the Source and partial Source complexes were the same, so that reaction rate constants for equations S4.17 – S4.18 and S4.21 – S4.22 were the same. Degradation of the single-stranded Initiator was likewise assumed to be governed by equations S4.25 – S4.27, which are analogous to equations S4.13 – S4.15 and were assumed to have the same rate constants fit for equations S4.13 and S4.15 in Section 4.5.5.2.

$$Source + Initiator \xrightarrow{kSInit} Output + Waste \tag{S4.16}$$

$$Source + Enzyme \xrightarrow{k_{ESf}} Source\colon Enzyme \tag{S4.17}$$

$$Source\colon Enzyme \xrightarrow{k_{ESrev}} Source + Enzyme \tag{S4.18}$$

$$Source\colon Enzyme \xrightarrow{k_{ESDPart}} Enzyme + Source_{Partial} \tag{S4.19}$$

$$Source_{Partial} + Initiator \xrightarrow{kS_{Part}Init} Output + Waste \tag{S4.20}$$

$$Source_{Partial} + Enzyme \xrightarrow{k_{ESf}} Source_{Partial}\colon Enzyme \tag{S4.21}$$

$$Source_{Partial}\colon Enzyme \xrightarrow{k_{ESrev}} Source_{Partial} + Enzyme \tag{S4.22}$$

$$Source_{Partial}\colon Enzyme \xrightarrow{k_{ESDeg}} Enzyme + Output \tag{S4.23}$$

$$Source_{Partial}\colon Enzyme \xrightarrow{k_{ESDeg}} Enzyme + Waste \tag{S4.24}$$

$$Initiator + Enzyme \xrightarrow{k_{EssDNAf}} Initiator\colon Enzyme \tag{S4.25}$$

$$Initiator\colon Enzyme \xrightarrow{k_{EInitrev}} Initiator + Enzyme \tag{S4.26}$$

$$Initiator\colon Enzyme \xrightarrow{k_{EssDNADeg}} Enzyme + Waste \tag{S4.27}$$

Finally, the degradation of the Source complex (rather than reaction with the Initiator) could release the Output which could then react with a Reporter complex. Equations S4.23 and S4.24 determine how much active Output strand is produced from degraded Source complexes. An active Output strand contains a contiguous to mostly contiguous toehold and domain (complementary to the Reporter) and is able to conduct strand-displacement reactions with the Reporter complex. Based upon the observation that Source complexes (mixed with Reporter) in the absence of Initiator strand showed a delayed increase in fluorescence (Figures S4.9 – S4.12), we modeled the degradation of Source complexes as a two-step process. Equations S4.2 – S4.15 were also included in the model, but were not fit, and the reaction rate constants for those reactions were set to their previously fit values (Tables S4.1 and S4.2). The degradation of Initiator strands was assumed to occur with the same model and reaction rate constants as Output strands (equations S4.25 – S4.27), except the Enzyme-Initiator unbinding reaction, which was fit individually for each Source complex.

Poorly synthesized or assembled Source complexes with 0 bp toeholds could in principle interact directly with the Reporter complex at rates comparable to their reaction with the Initiator, as observed in similar reactions in nuclease-free buffers.[101] Thus, our model of 0 bp strand-displacement between the Source and Initiator also included two additional reactions to reflect this possibility.[101]

$$Source_{Leak} + Reporter \xrightarrow{kS_{Leak}R} Fluorophore + Quencher \qquad (S4.28)$$

$$Source_{Leak} + Initiator \xrightarrow{kS_{Leak}I} Output + Waste \qquad (S4.29)$$

where $Source_{Leak}$ represents Source complexes that are partially formed or have mismatches, due to synthesis or annealing errors, that enable it to react with Reporter complexes in the absence of Initiator strands. The reaction rate constants $kS_{Leak}R$ and $kS_{Leak}I$ were assumed to be 5.5×10^3 1/M-sec and 5 1/M-sec which represent ~4 bp and 1 bp toeholds, respectively.[101] The concentration of $\text{Source}_{\text{Leak}}$ was assumed to be 4% of the total initial Source concentration, Equation S4.28 only influences the level of fluorescence/Output detected by the Reporter prior to the addition of Initiator or prior to nucleases beginning to significantly degrade Source complexes (*e.g.*, within 6 hours of Source addition to the reaction mixture). Equation S4.29 directs the initial slope of the curves immediately after Initiator is added because Initiator presumably reacts with $\text{Source}_{\text{Leak}}$ 10-fold faster than with Source complexes. The reactions between $\text{Source}_{\text{Leak}}$ complexes and nucleases were not included in our model because the $\text{Source}_{\text{Leak}}$ complexes were assumed to be depleted *via* the reactions shown in equations S4.28 and S4.29 before significant degradation of $\text{Source}_{\text{Leak}}$ complexes occurs.

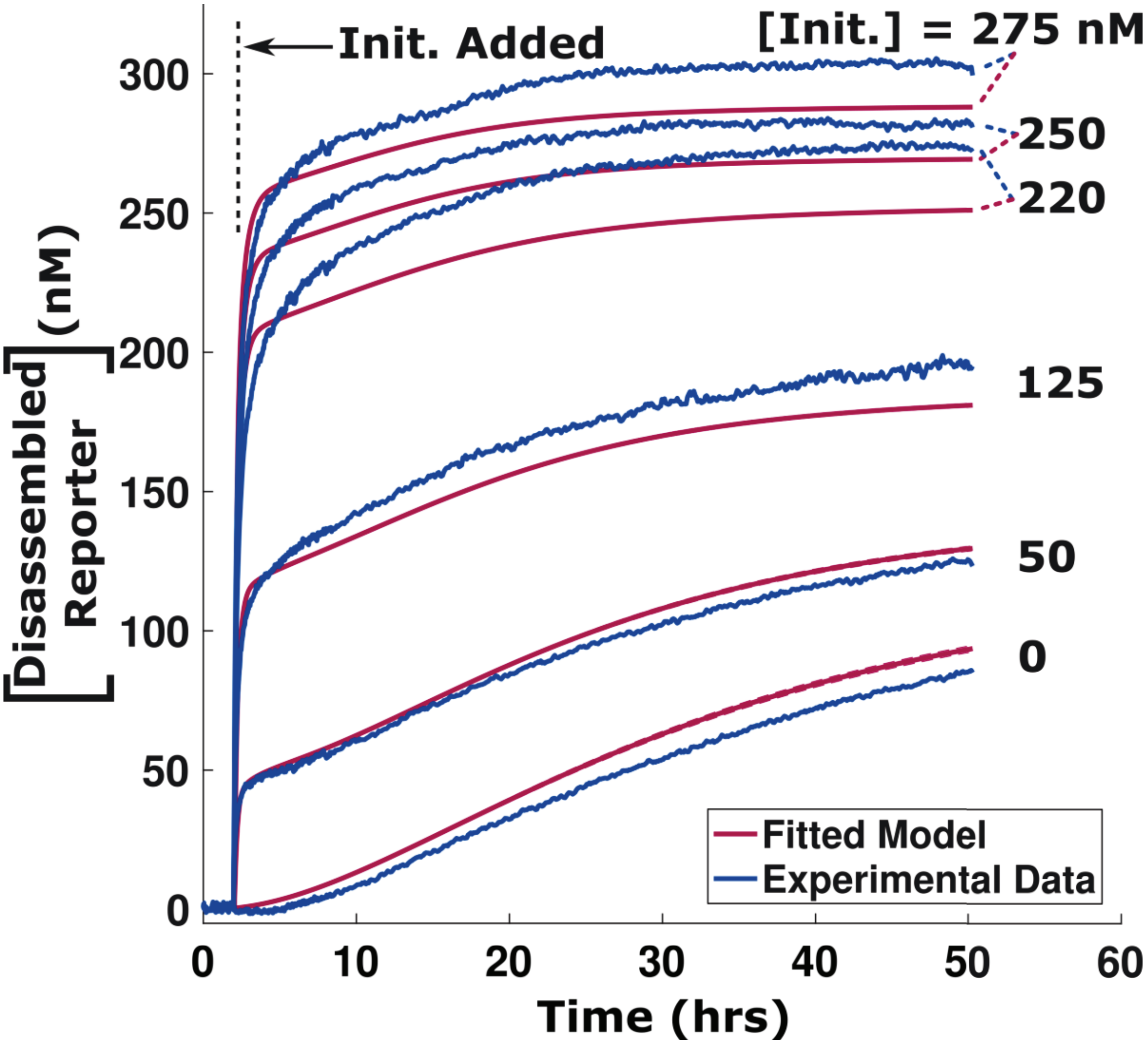


Figure S4.9: Model predictions and experimental data for a system in which the Source complex can react with an Initiator strand using a 5 base pair toehold to release an Output strand. The Output strand can displace the quencher-containing strand from the Reporter complex to increase the measured fluorescence. The process and resulting parameters, including those that govern the speed of 5 base pair toehold-mediated strand-displacement in nuclease-screened medium, are given in Section 4.5.5.3 and Table S4.3. The model kinetics are predicted [Disassembled Reporter] values. Confidence intervals (95%) for the fit are too small to be observed.

Table S4.3: List of fit values for the parameters for our model of strand-displacement of a Source complex by an Initiator given a 5 base pair toehold (see Section 4.5.5.3). Data used for fitting and results are shown in Figure S4.9. Parameters are listed as the fitted value +/- their 95% confidence interval in parentheses.

Reaction Component	**Expt. Conc. (nM)**	**Modeled Conc. (nM)**	**Rate Constant**	**Fitted Parameter**
Reporter	200	200, 200, 220, 280, 290, 310	k_{SInit} (1/M-sec)	7(1)x10^{4}
Initiator	0, 50, 100, 150, 200, 300	0, 50, 125, 220, 250, 275	$k_{SpartInit}$ (1/M-sec)	4(2)x10^{4}
Source (5bp)	200	275	k_{ESf} (1/M-sec)	4(2)x10^{4}
Enzyme		31.3	k_{ESrev} (1/sec)	2.2(0.3)x10^{-1}
Inhibitor		620	k_{ESDeg} (1/sec)	4(3)x10^{-2}
			$k_{EInitrev}$ (1/sec)	3(4)x10^{-2}

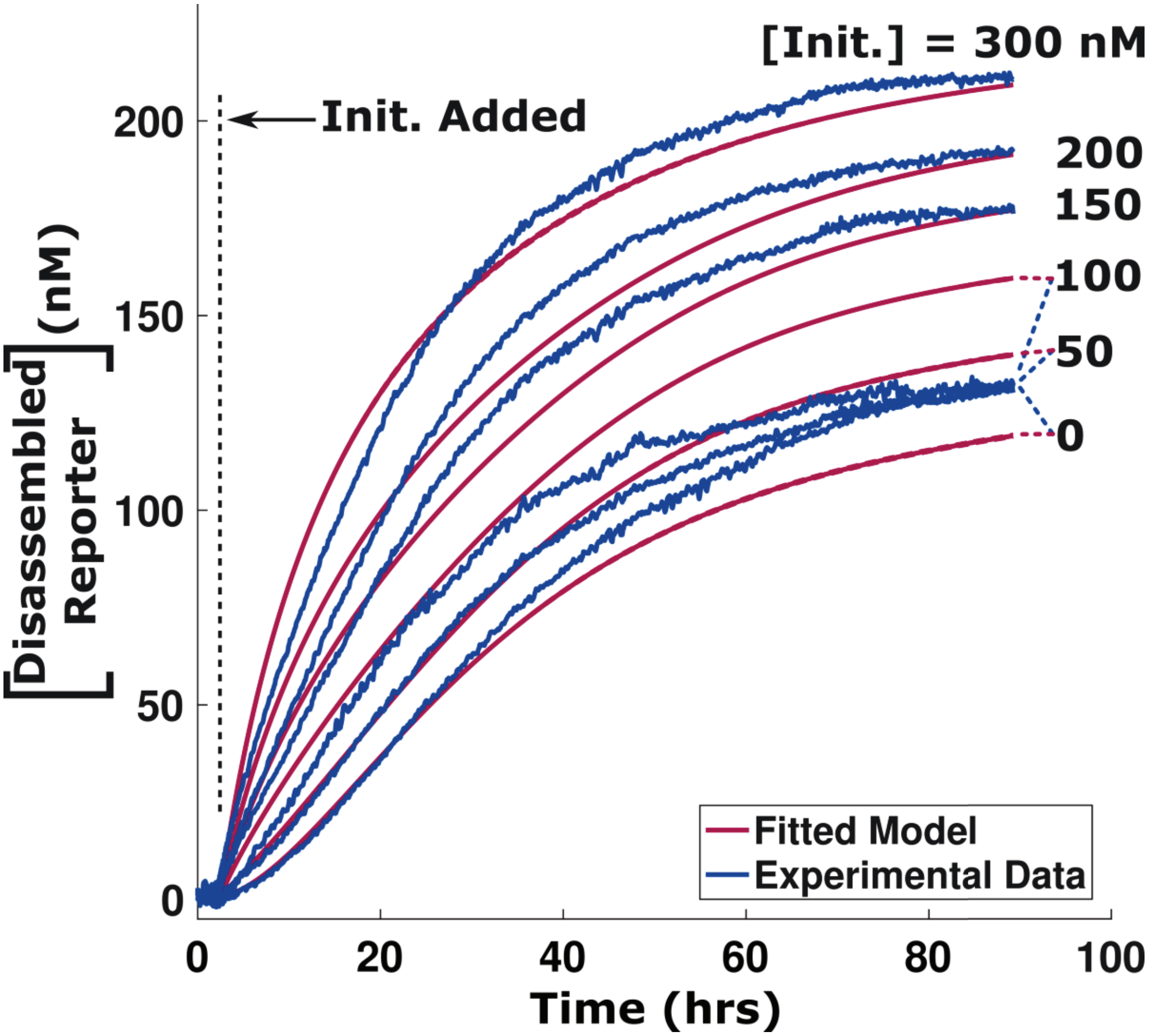


Figure S4.10: Model predictions and experimental data for a system in which the Source complex can react with an Initiator strand using a 2 base pair toehold to release an Output strand. The Output strand can displace the quencher-containing strand from the Reporter complex to increase the measured fluorescence. The fitting process and resulting parameters, including those that govern the speed of 2 base pair toehold-mediated strand-displacement in nuclease-screened medium, are given in Section 4.5.5.3 and Table S4.4. The model kinetics are predicted [Disassembled Reporter] values. Confidence intervals (95%) for the fit are too small to be observed.

Table S4.4: List of fit values for the parameters for our model of strand-displacement of a Source complex by an Initiator given a 2 base pair toehold (see Section 4.5.5.3). Data used for fitting and results are shown in Figure S4.10. Parameters are listed as the fitted value +/- their 95% confidence interval in parentheses.

Reaction Component	**Expt. Conc. (nM)**	**Modeled Conc. (nM)**	**Rate Constant**	**Fitted Parameter**
Reporter	200	225	k_{SInit} (1/M-sec)	$6.5(0.3)\text{x}10^{1}$
Initiator	0, 50, 100, 150, 200, 300	0, 50, 100, 150, 200, 300	$k_{SpartInit}$ (1/M-sec)	$1.5(0.5)\text{x}10^{2}$
Source (2bp)	200	220	k_{ESf} (1/M-sec)	$1.5(1.2)\text{x}10^{4}$
Enzyme		31.3	k_{ESrev} (1/sec)	$2.1(1.5)\text{x}10^{-2}$
Inhibitor		620	k_{ESDeg} (1/sec)	$2.6(4.8)\text{x}10^{-2}$
			$k_{EInitrev}$ (1/sec)	$5.1(0.7)\text{x}10^{-2}$

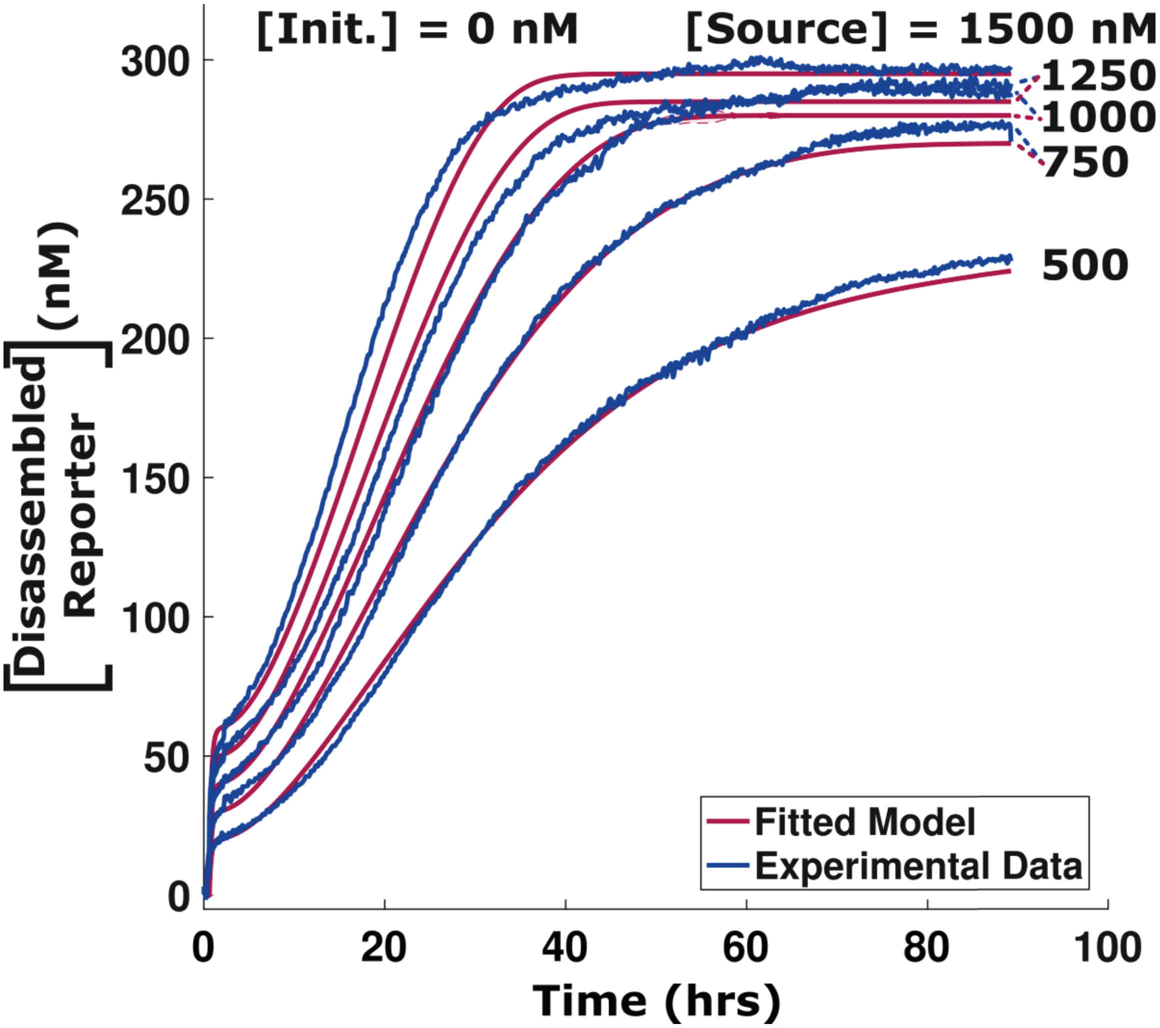


Figure S4.11: Model predictions and experimental data for a system in which the 0 bp toehold Source complex can be digested by nucleases in the absence of Initiator to release an Output strand. The Output strand can displace the quencher-containing strand from the Reporter complex to change the fluorescence. The process and resulting parameters, including those that govern the speed of 0 base pair toehold-mediated strand-displacement in nuclease-screened media, are given in Section 4.5.5.3 and Table S4.5. The model kinetics are predicted [Disassembled Reporter] values. Confidence intervals (95%) for the fit are too small to be observed.

Table S4.5: List of fit values for the parameters for our model of nuclease-0 bp toehold Source complex reactions in the absence of Initiator (see Section 4.5.5.3). Reaction components with multiple modeling concentrations are listed in order in regards to the other components (*e.g.*, [Reporter]=270 nM and [Source]=500 nM are initial concentrations for one set of modeled reaction mixtures). Data used for fitting and results are shown in Figure S4.11. Parameters are listed as the fitted value +/- their 95% confidence interval in parentheses.

Reaction Component	**Expt. Conc. (nM)**	**Modeled Conc. (nM)**	**Rate Constant**	**Fitted Parameter**
Reporter	200	270, 270, 280, 285, 295	k_{ESf} (1/M-sec)	1.2(0.3)x10^{4}
Initiator	0	0	k_{ESrev} (1/sec)	5(100)x10^{-5}
Source	500, 750, 1000, 1250, 1500	500, 750, 1000, 1250, 1500	k_{ESDeg} (1/sec)	3.1(0.2)x10^{-3}
Enzyme		31.3		
Inhibitor		620		

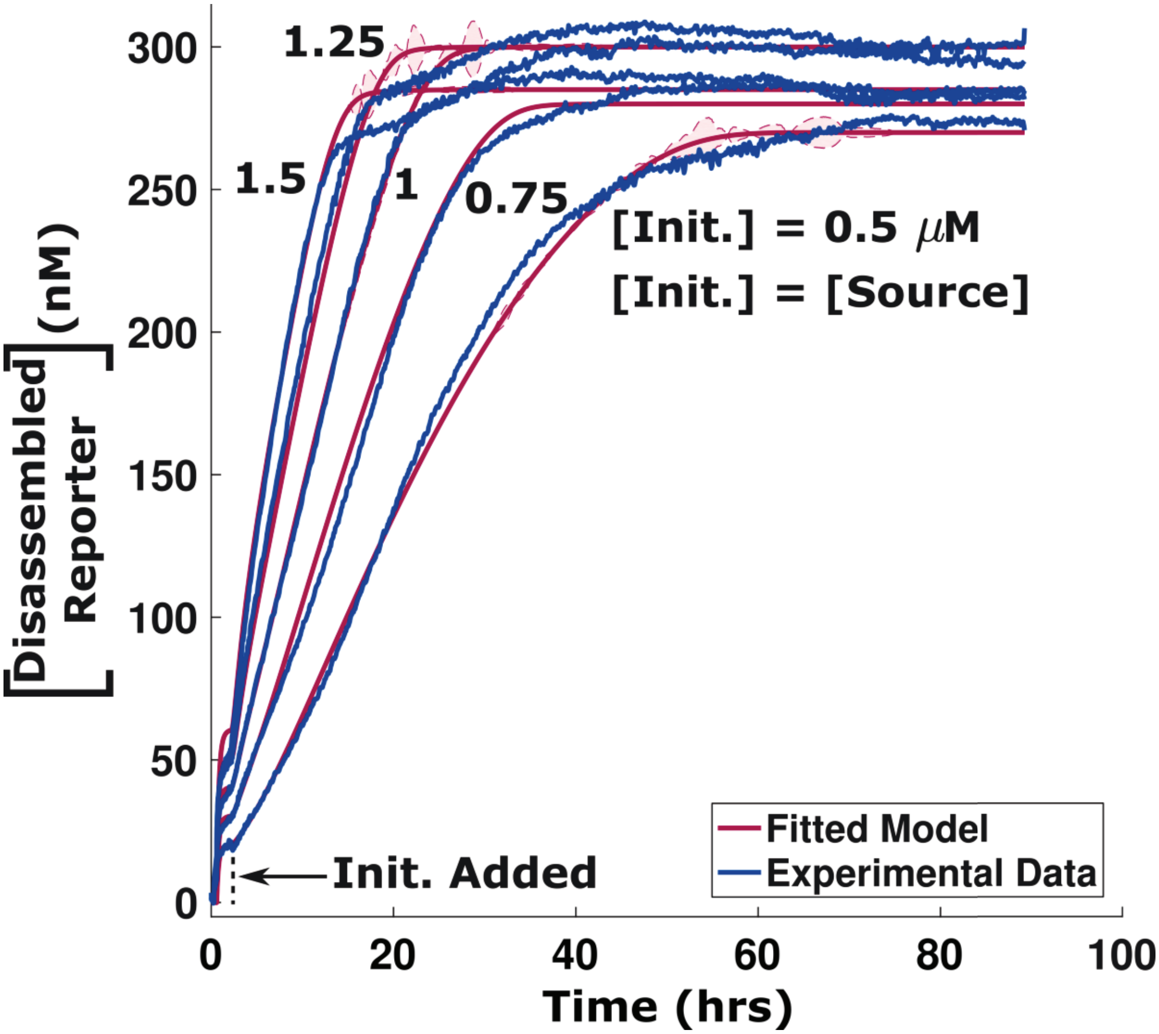


Figure S4.12: Model predictions and experimental data for a system in which the Source complex can react with an Initiator strand using a 0 base pair toehold to release an Output strand. The Output strand can displace the quencher-containing strand from the Reporter complex to increase the measured fluorescence. The process and resulting parameters, including those that govern the speed of 0 base pair toehold-mediated strand-displacement in nuclease-screened medium, are given in Section 4.5.5.3 and Table S4.6. The model kinetics are predicted [Disassembled Reporter] values. Confidence intervals (95%) for the fit are too small to be observed.

Table S4.6: List of fit values for the parameters for our model of strand-displacement of a Source complex by an Initiator given a 0 base pair toehold (see Section 4.5.5.3). Reaction components with multiple modeling concentrations are listed in order in regards to the other components (*e.g.*, [Reporter]=270 nM and [Source]=500 nM are initial concentrations for one set of modeled reaction mixtures). Data used for fitting and results are shown in Figure S4.12. Parameters are listed as the fitted value +/- their 95% confidence interval in parentheses.

Reaction Component	**Expt. Conc. (nM)**	**Modeled Conc. (nM)**	**Rate Constant**	**Fitted Parameter**
Reporter	200	270, 280, 300, 300, 285	k_{SInit} (1/M-sec)	$9(1)x10^{-1}$
Initiator	500, 750, 1000, 1250, 1500	500, 750, 1000, 1250, 1500	$k_{SpartInit}$ (1/M-sec)	$9.3(0.7)x10^{1}$
Source	500, 750, 1000, 1250, 1500	500, 750, 1000, 1250, 1500	k_{ESf} (1/M-sec)	$6(6)x10^{3}$
Enzyme		31.3	k_{ESrev} (1/sec)	$2(200)x10^{-4}$
Inhibitor		620	k_{ESDeg} (1/sec)	$2(3)x10^{-2}$
			$k_{EInitrev}$ (1/sec)	1.3(0.1)

4.5.6 Predicting the behavior of complex DNA strand-displacement circuits

Complex strand-displacement circuits could make it possible to integrate information about the concentrations of different molecules and direct multi-faceted responses with controlled responses in serum. To design such complex circuits, tools for predicting their behavior in advance will be needed. In this section, we use our experimental characterization of DNA strand-displacement reactions to predict the behavior of multi-stage reaction cascades and a timer to work toward building models of these systems, and to determine what changes will be needed in order to build reliably robust circuits of these types. The models we construct build on the model introduced in Section 4.5.5.

4.5.6.1 Multi-layer cascade circuit simulations

To model the kinetics of a strand-displacement cascade with multiple Sources and Initiators (Figure 4.6), we expanded the model presented in Section 4.5.5 with strand-displacement reactions between the additional components and reactions that modeled their degradation. Each layer of the cascade is modeled by a set of reactions like those for the first layer, *i.e.*, equations S4.16 – S4.27. Additional reactions allow partially degraded Source complexes to release Initiator strands that can signal the next layer of the cascade, analogous to the reaction shown in equation S4.23. The resulting set of equations involving either the Source or Initiator or their partially degraded versions of these species in an n-layer circuit are therefore:

$$Source_n + Initiator_n \xrightarrow{kSInit} Initiator_{n-1} + Waste \tag{S4.30}$$

$$\begin{array}{l} \vdots \\ \vdots\; n - circuit\ layers \\ \vdots \end{array}$$

$$Source_1 + Initiator_1 \xrightarrow{kSInit} Output + Waste \tag{S4.31}$$

$$Source_n + Enzyme \xrightarrow{k_{ESf}} Source_n\!: Enzyme \tag{S4.32}$$

$$Source_n\!: Enzyme \xrightarrow{k_{ESrev}} Source_n + Enzyme \tag{S4.33}$$

$$Source_n\!: Enzyme \xrightarrow{k_{ESDPart}} Enzyme + Source_{n_{Partial}} \tag{S4.34}$$

$$Source_{n_{Partial}} + Initiator \xrightarrow{kS_{Part}Init} Initiator_{n-1} + Waste \tag{S4.35}$$

$$Source_{n_{Partial}} + Enzyme \xrightarrow{k_{ESf}} Source_{n_{Partial}}\!: Enzyme \tag{S4.36}$$

$$Source_{n_{Partial}}\!: Enzyme \xrightarrow{k_{ESrev}} Source_{n_{Partial}} + Enzyme \tag{S4.37}$$

$$Source_{n_{Partial}}: Enzyme \xrightarrow{k_{ESDeg}} Enzyme + Initiator_{n-1} \qquad (S4.38)$$

$$Source_{n_{Partial}}: Enzyme \xrightarrow{k_{ESDeg}} Enzyme + Waste \qquad (S4.39)$$

where *n* indicates the layer number of the cascade circuit. For example, in a 2-layer cascade, Initiator$_2$ and Source$_2$ react to release Initiator$_1$ (equation S4.30). Initiator$_1$ can then react with Source$_1$ to release the Output strand (equation S4.31). Parameters for the above reactions were taken from Tables S4.1 – S4.6 above. Degradation and DNA strand-displacement reaction rate constants for Source complexes were chosen based upon the toehold length (*e.g.*, Source with a 5 bp toehold was simulated with parameters from Table S4.3). As described in Section 4.5.5, there was a discrepancy between the simulated concentrations and the experimental concentrations of reaction components that generated the best fit to the data. Along the lines of generating the best prediction for experimentally derived data, we simulated the cascade system with 250 nM of Reporter complexes, which may be assumed to experimentally match a Reporter concentration of 200 nM as demonstrated in Section 4.5.5. Source complexes were assumed to have matching experimental and simulation concentrations since the variance was only detected in the 2 bp and 5 bp toehold Release reaction cases shown in Section 4.5.5.

Table S4.7: List of concentrations used for simulating the reaction between Initiator and Source complexes in the 2-layer cascade reaction systems incubated in nuclease-screened medium.

Reaction Component	Modeled Conc. (nM)	Expected Expt. Conc. (nM)
Reporter	250	200
Initiator$_2$	0, 50, 100, 150, 200, 300	0, 50, 100, 150, 200, 300
Source$_1$ = Source$_2$	200	200

We used this model to predict the kinetics of a 2-layer strand displacement cascade with toehold lengths of 5 bp or 2 bp for both layers, as well a cascade with 2 layers with toehold lengths of 5 and 2 bp for the 2nd and 1st layer of the circuit. To understand the kinetics of the cascade, we began by simulating these two-level cascade circuits in the absence of nucleases (*e.g.*, serum-free medium) in order to compare the effect of toehold length in the multi-layer circuits (Figure S4.13). In this case, all nuclease-dependent reactions are omitted from the model and the reaction rate constants for the remaining reactions were taken from Tables S4.1 – S4.6. In this serum-free case, the varying input Initiator concentrations were clearly distinguishable from one another.

Simulation of the 2-layer cascade in the presence of nuclease enzymes showed that adding multiple layers of circuits to a system compounded the effect of nuclease degradation on the output values (Figures S4.14 – S4.16). In all cases, a reduced conversion of input Initiator_2 strand concentration into detected Output concentration was observed, *i.e.*, there was a narrower range of initial [Initiator_2] that produced different levels of output fluorescence with a two-layer cascade than with a 1-layer cascade.

The model was able to predict the reduction in input:output conversion for 2-layer cascades with 2 bp for both layers or 5 bp (1st layer)/2 bp (2nd layer) toeholds observed in experiments, but experiments of 2-layer cascades with only 5 bp toeholds showed an even greater loss of input-to-output conversion than our model predicted (Figure S4.14). Overall, the simulations suggest that circuit robustness decreases significantly with circuit depth, indicating that better methods of preventing degradation are needed to reliably operate more complex circuits, such as Boolean logic circuits.[87,107]

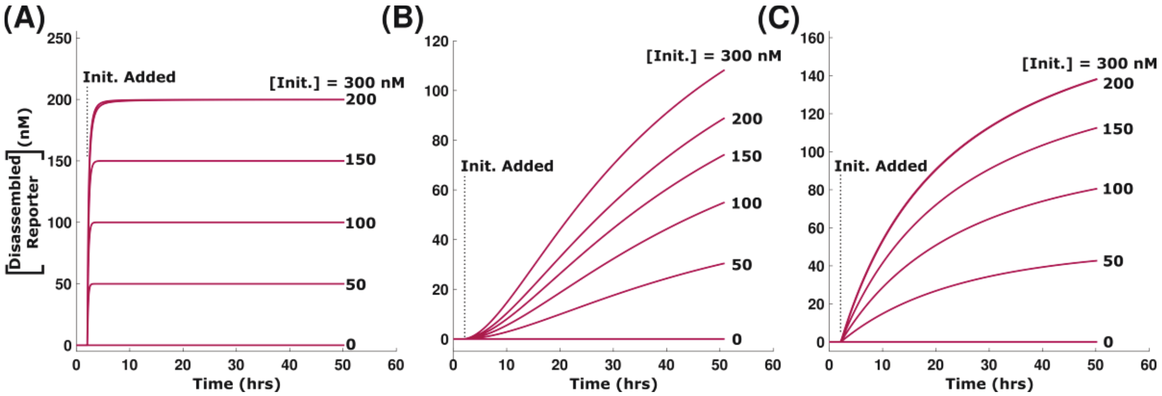


Figure S4.13: Simulation of a 2-layer cascade circuit in nuclease-free conditions (see Figure 4.6A). (A) Simulated kinetics of a 2-layer cascade with 5 bp toeholds on both Source complexes in the cascade. (B) Simulated kinetics of a 2-layer cascade with 2 bp toeholds on both Source complexes in the cascade. (C) Simulated kinetics of a 2-layer cascade with a Source complex with a 5 bp toehold on the first layer and a 2 bp toehold on the second layer. Concentrations of reaction components are listed in Table S4.7, except that all enzyme concentrations were set to zero. Reaction rate constants are listed in Tables S4.1 – S4.6.

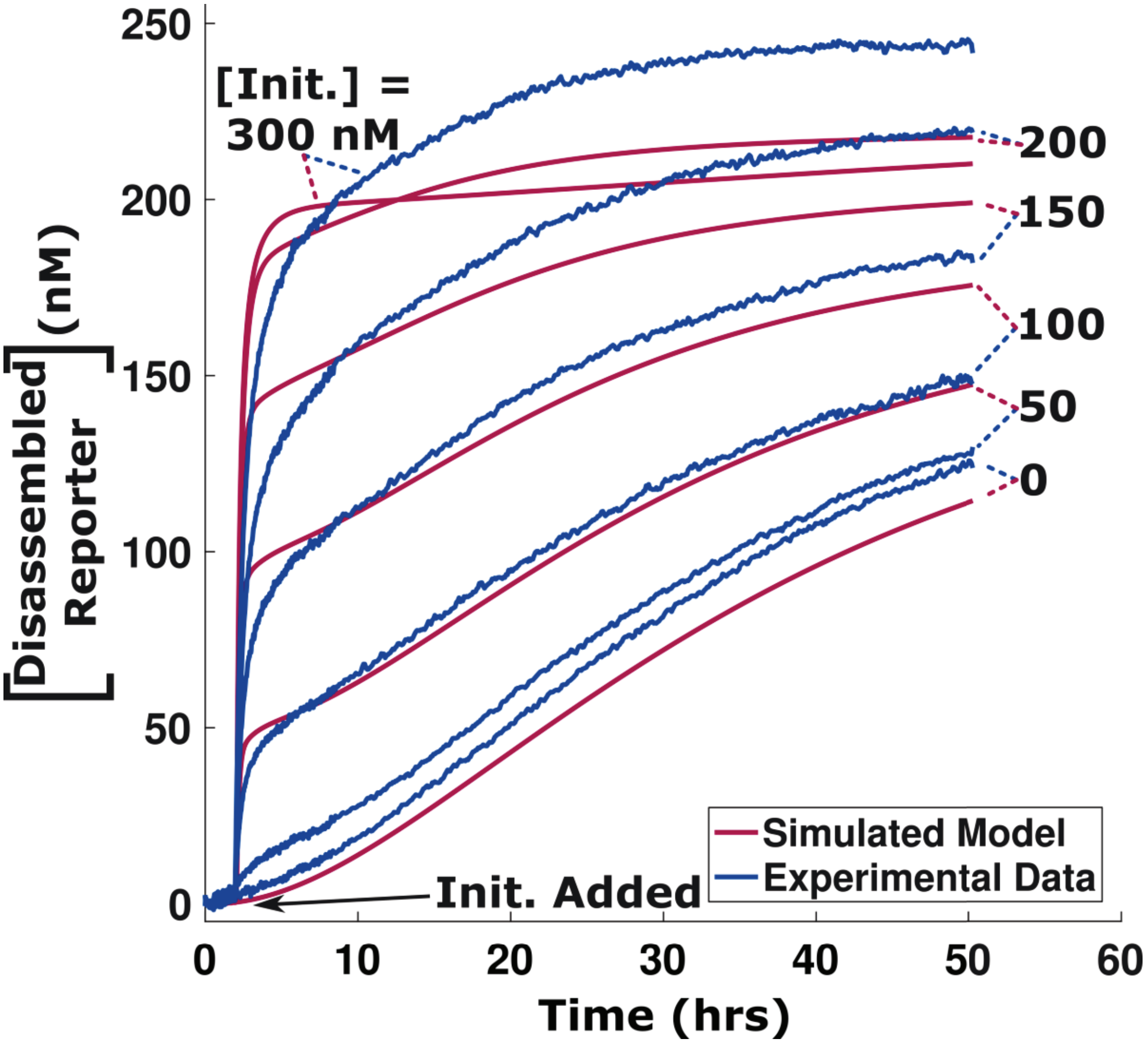


Figure S4.14: Simulation of a 2-layer cascade with 5 bp toeholds on both Source complexes in the cascade in nuclease-screened medium. The concentrations of reaction components are listed in Table S4.7. Experiments showed a slower initial rise of Disassembled Reporter signal upon addition of Initiator$_2$, suggesting less release of the final Output strand than experiments. There was also a lower range of initial [Initiator$_2$] that could be distinguished by final fluorescence than the model predicts.

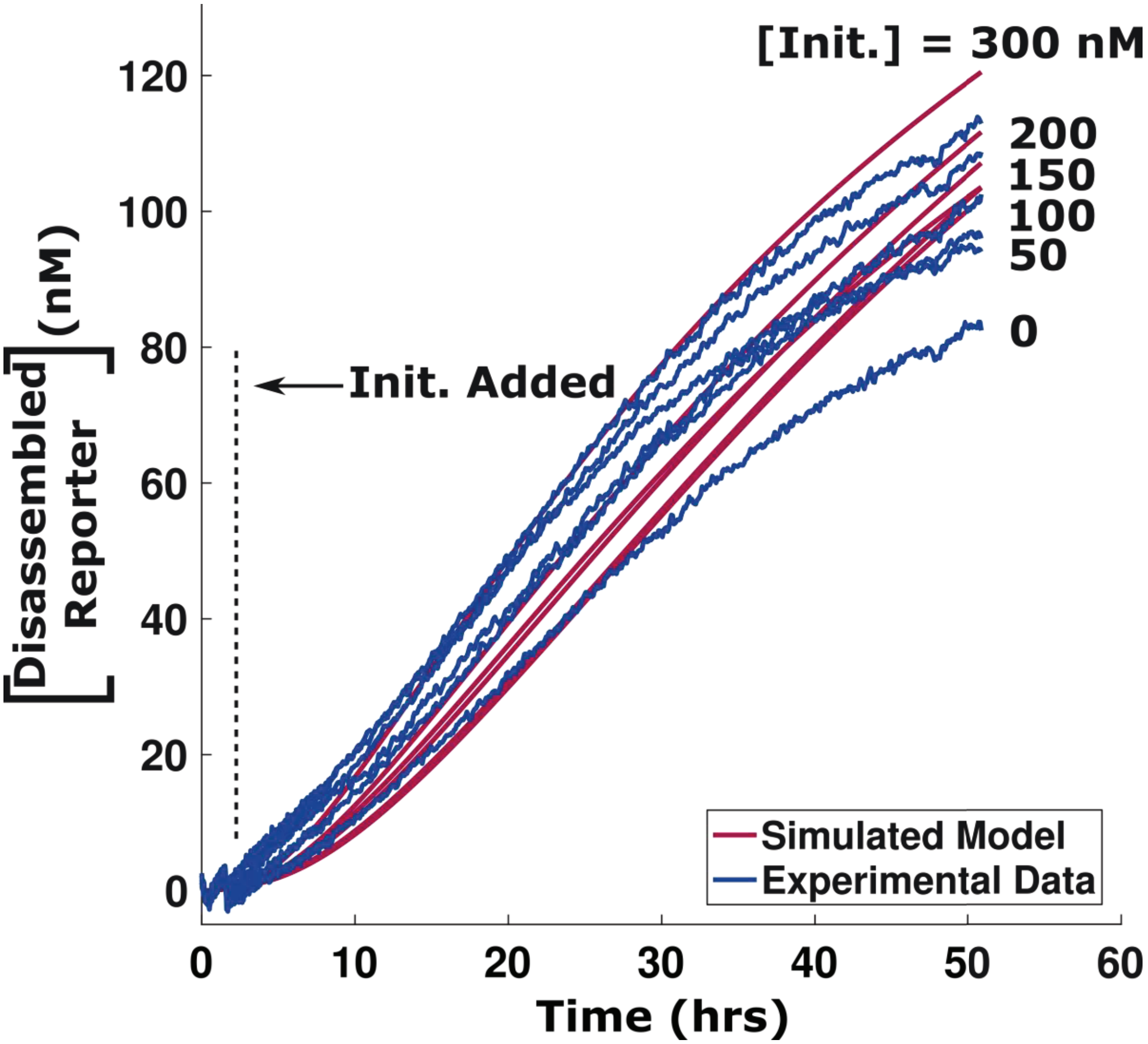


Figure S4.15: Simulation of a 2-layer cascade with 2 bp toeholds on both Source complexes in the cascade in nuclease-screened medium. The concentrations of reaction components are listed in Table S4.7. The simulations showed comparable Output strand release rates to what was observed in experiment. In general, it appeared that the release rate of the Output of the 2-layer cascade in this circuit, as measured by interaction with the Reporter complex, was dominated by degradation processes mediated by nucleases rather than by strand-displacement.

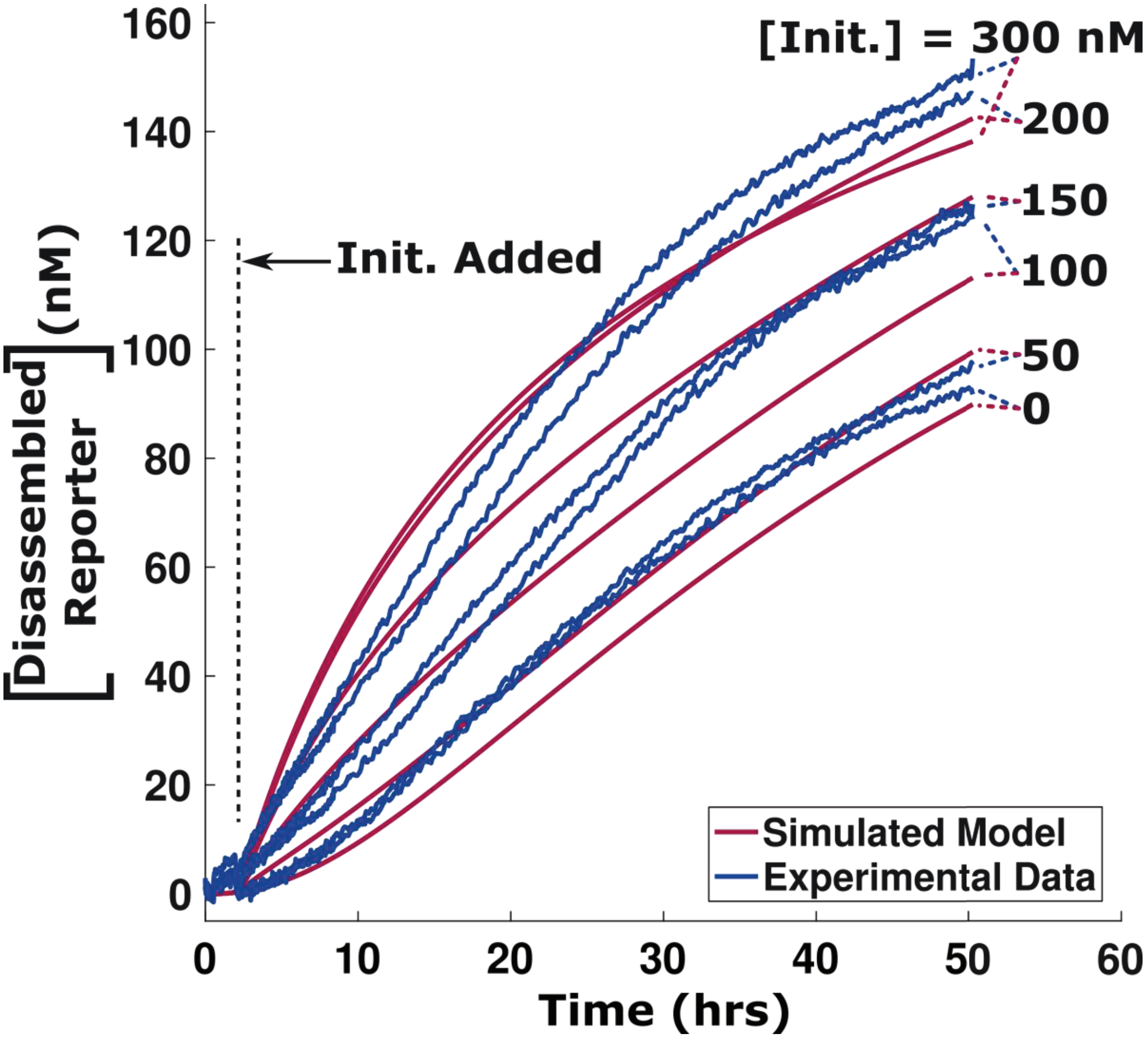


Figure S4.16: Simulation of a 2-layer cascade with 5 bp toeholds on the first Source complex and 2 bp toeholds on the second Source complex in nuclease-screened medium. The concentrations of reaction components are listed in Table S4.7. The simulations showed comparable Output strand release rates to what was observed in experiment. As in the 2-layer cascade in which both Source complexes had 2 bp toeholds, the release of Output appears to be largely directed by nuclease-mediated degradation.

4.5.6.2 Timer circuits in nuclease-screened medium

To understand more about how circuits for controlled release might operate in serum-supplemented medium, we used a timer circuit which uses a slow release step coupled to a fast recapturing process to delay the overall production of the output species until the desired time.[101] When another process (*e.g.*, reporting, directing cellular behavior, or nanostructure assembly) is added downstream of the timer circuit, there is a competition between the downstream process and the recapturing process for the output species being slowly released. Thus, it is of interest to understand how degradation processes across all species influence the timing and competitive processes occurring within the timer circuit coupled to a downstream Reporter (Figure S4.17), and to determine how well control over the timing of release can be achieved in serum.

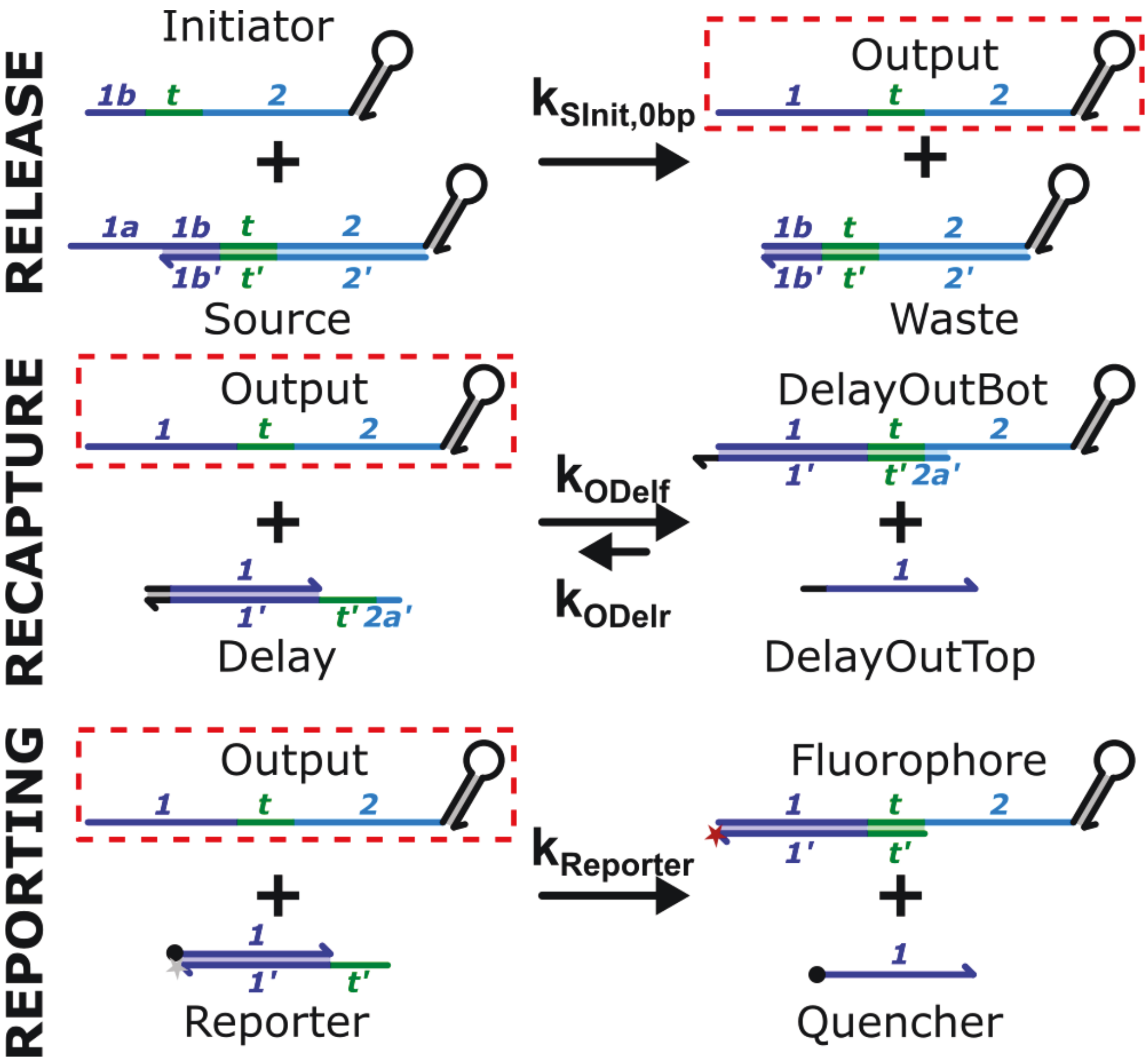


Figure S4.17: Schematic of a timer circuit coupled to a downstream reporting process. The Output strand (red boxes) is slowly released from the Source complex using a 0 bp toehold initiation process. The Output can either be quickly recaptured using a Delay complex with a 7 bp toehold, or detected using a Reporter complex with a 5 bp toehold. The recapture process has a forward reaction rate constant ~20-fold larger than the reporting process.

To simulate the kinetics of the timer circuit in serum-supplemented medium, we constructed a model based on the ideas in Section 4.5.5, beginning with reaction equations S4.2 – S4.29. Reactions were then added to account for the added DNA circuit species, *Delay*, which was also assumed to be degraded by nucleases. Additionally, the following undesired "leak" reactions between the DNA components of the timer were also incorporated into the model as previously described[101] and as follows:

$$Output + Delay \xrightarrow{kODelf} DelOutTop + DelOutBot \tag{S4.40}$$

$$DelOutTop + DelOutBot \xrightarrow{kODelr} Output + Delay \tag{S4.41}$$

$$Source_{Leak} + Delay \xrightarrow{kS_{Leak}Del} OutDelBot_{Leak} + DelOutTop \tag{S4.42}$$

$$Source + Delay \xrightarrow{kSDel} DelOutTop + DelOutBot \tag{S4.43}$$

$$Delay + Enzyme \xrightarrow{kEDelf} Enzyme{:}\, Delay \tag{S4.44}$$

$$Enzyme{:}\, Delay \xrightarrow{k_{EDelr}} Enzyme + Delay \tag{S4.45}$$

$$Enzyme{:}\, Delay \xrightarrow{kEDelDeg} Enzyme + Waste \tag{S4.46}$$

$$DelOutBot + Enzyme \xrightarrow{kEDelf} Enzyme{:}\, DelOutBot \tag{S4.47}$$

$$Enzyme{:}\, DelOutBot \xrightarrow{k_{EDelr}} Enzyme + DelOutBot \tag{S4.48}$$

$$Enzyme{:}\, DelOutBot \xrightarrow{k_{EDelDeg}} Enzyme + Output \tag{S4.49}$$

$$Enzyme{:}\, DelOutBot \xrightarrow{k_{EDelDeg}} Enzyme + Waste \tag{S4.50}$$

$$Initiator + Delay \xrightarrow{kInitDelf} Init{:}\, Del \tag{S4.51}$$

$$Init{:}\, Del \xrightarrow{kInitDelr} Initiator + Delay \tag{S4.52}$$

where *DelOutTop* and *DelOutBot* correspond to the top and bottom complex of the Delay/Output reaction. Reaction rate constants of the Delay-enzyme and DelOutBot-enzyme reactions were chosen to be the same as Reporter-enzyme reactions (Section 4.5.5.1). Equations S4.51 and S4.52 represent the reversible binding between Initiator and Delay complexes with reaction rate constants calculated as in Zhang and Winfree.[96] The off-rate k_{InitDelr} was calculated using "Nupack, dangles=some" parameters for a 7 bp toehold at 37 °C. The reaction rate constants in equations S4.40 – S4.43 were chosen as previously described[101] with minor adjustments made for reactions being conducted at 37 °C and in serum-supplemented medium (Table S4.8). Reaction rate constants for Source degradation are listed in Table S4.5 (no Initiator) and Table S4.6 (with Initiator). Simulated component concentrations are listed in Table S4.9.

Table S4.8: Reaction rate constants for the simulation of the timer circuit as described in equations S4.40 – S4.52.

Rate Constant	Parameter Value
k_{ODelf} (1/M-sec)	5x10^{5}
k_{ODelr} (1/M-sec)	5x10^{1}
$k_{SLeakDel}$ (1/M-sec)	2.5x10^{4}
k_{SDel} (1/M-sec)	2.5
$k_{InitDelf}$ (1/M-sec)	3.5x10^{6}
$k_{InitDelr}$ (1/sec)	0.291

Table S4.9: List of parameters used for simulating the reaction between Initiator and Source complexes in the timer circuit system incubated in serum-supplemented medium. The concentration of Reporter in the simulation was higher than the expected experimental concentration because the measured, calibrated concentration of Disassembled Reporter usually increased beyond the stoichiometric limit of the initial Reporter concentration in the experimental reaction mixture (Section 4.5.4).

Reaction Component	**Modeled Conc. (nM)**	**Expected Expt. Conc. (nM)**
Reporter	250	200
Initiator	0 or 750	0 or 750
Source	750	750
Delay	0, 40, 80, 120, 160	0, 40, 80, 120, 160

As with the cascade circuit in Section 4.5.6.1, we characterized the effects of nuclease enzymes on the system by using the model to predict kinetics for the timer circuit with all enzyme concentrations set to zero (Figure S4.18). Because downstream processes can compete with the Delay complex for released Output, we simulated the enzyme-free case in the presence of and without Reporter. When an irreversible Reporter is used to detect the release of Output from the timer circuit, it competes with the Delay complex for released Output strands to "load" the circuit and prevent the desired delayed-release behavior observed in the absence of Reporter complexes. This is in contrast to the reversible reporting used previously[101] that enabled Delay complexes to more favorably compete for released Output and kept the detected [Disassembled Reporter] low until the concentration of Delay complexes was effectively zero. Since reversible reporting requires continuous strand-displacement exchanges that keeps either a quencher-modified strand or the Output strand in a single-stranded form, the reporting process is more susceptible nuclease-mediated degradation that would disturb the process over time as an increasing amount of those single-stranded components are degraded.

In the presence of nucleases, Source complexes can degrade and release "active" Output molecules in the absence of Initiator, thus we considered both initiated and un-initiated cases (Figures S4.19 – S4.20). The timer circuit simulations show that while the model can reliably predict the behavior of the circuit, including the addition of an experimentally un-tested reaction component (Delay complex), the operability of the circuit is severely diminished due to the presence of the nucleases, again indicating the need for a more robust protection method. Experimental measurements of the system showed the same trends, but with an apparently lower Source degradation rate. This could be due to Delay complexes having a lower degradation rate than what was assumed.

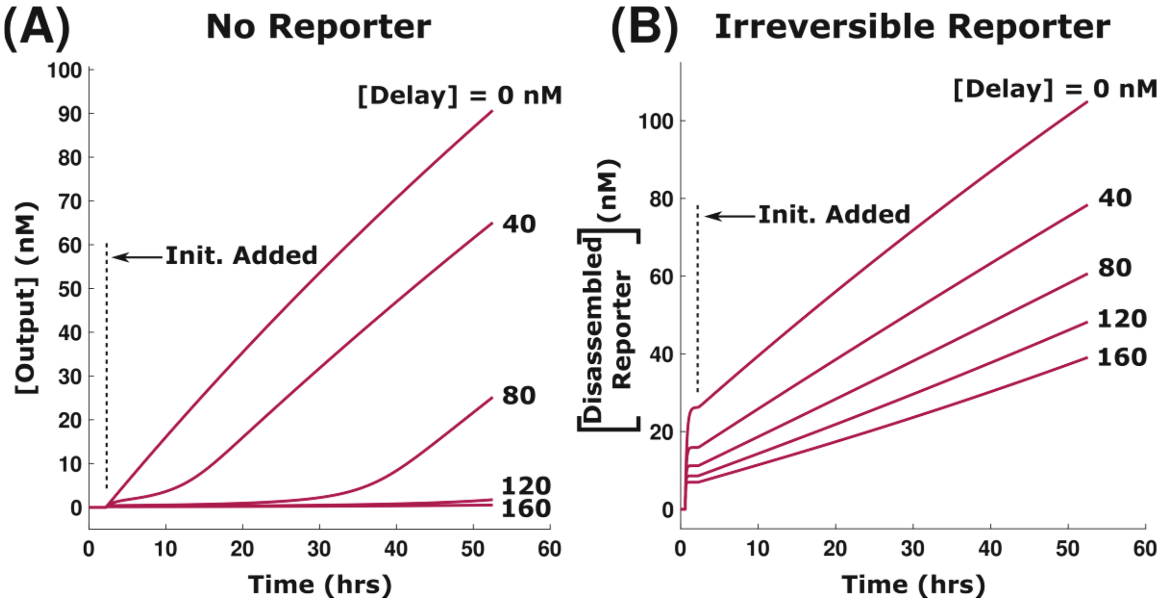


Figure S4.18: Simulations of the release of Output in the timer circuit (Figure S4.17) with all enzyme concentrations set to zero in the model. (A) Simulated kinetics of release of Output from the timer circuit without the Reporter complex. (B) Simulated kinetics of release of the Output from timer measured using the irreversible Reporter. Concentrations of reaction components are listed in Table S4.9, except that the concentration of all enzymes are set to zero. The Reporter competes with the Delay complex for released Output, loading the circuit, and preventing the delayed release behavior observed in (A).

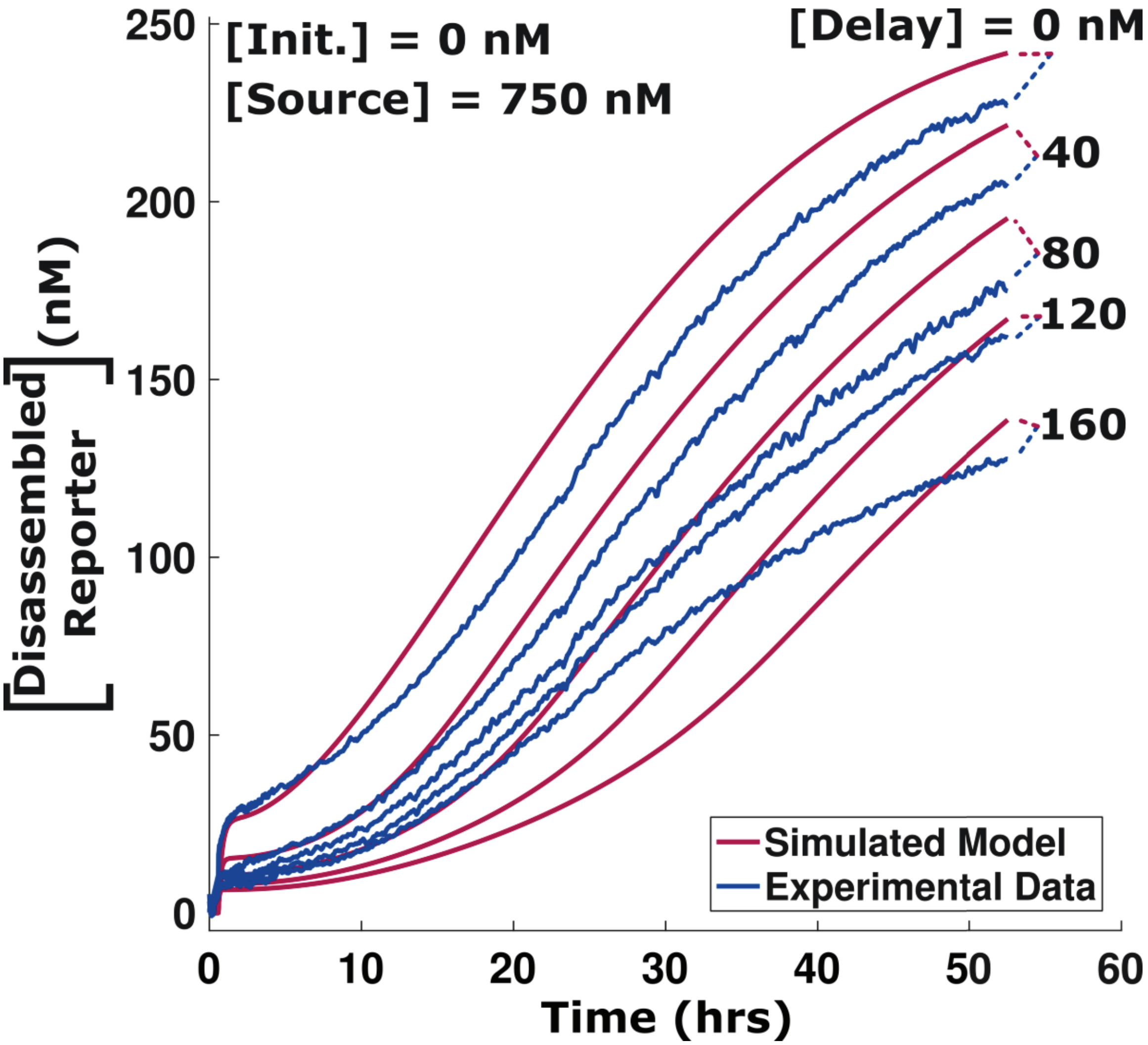


Figure S4.19: Release of the timer circuit Output (as measured by its interaction with the Reporter complex) in the absence of Initiator. In this case, release is either a consequence of "leak" reactions, or degradation of the strand-displacement reaction components. Concentrations of reaction components are listed in Table S4.9. Experiments and simulations show similar trends, although there is a longer lag time to release in simulations than in experiments.

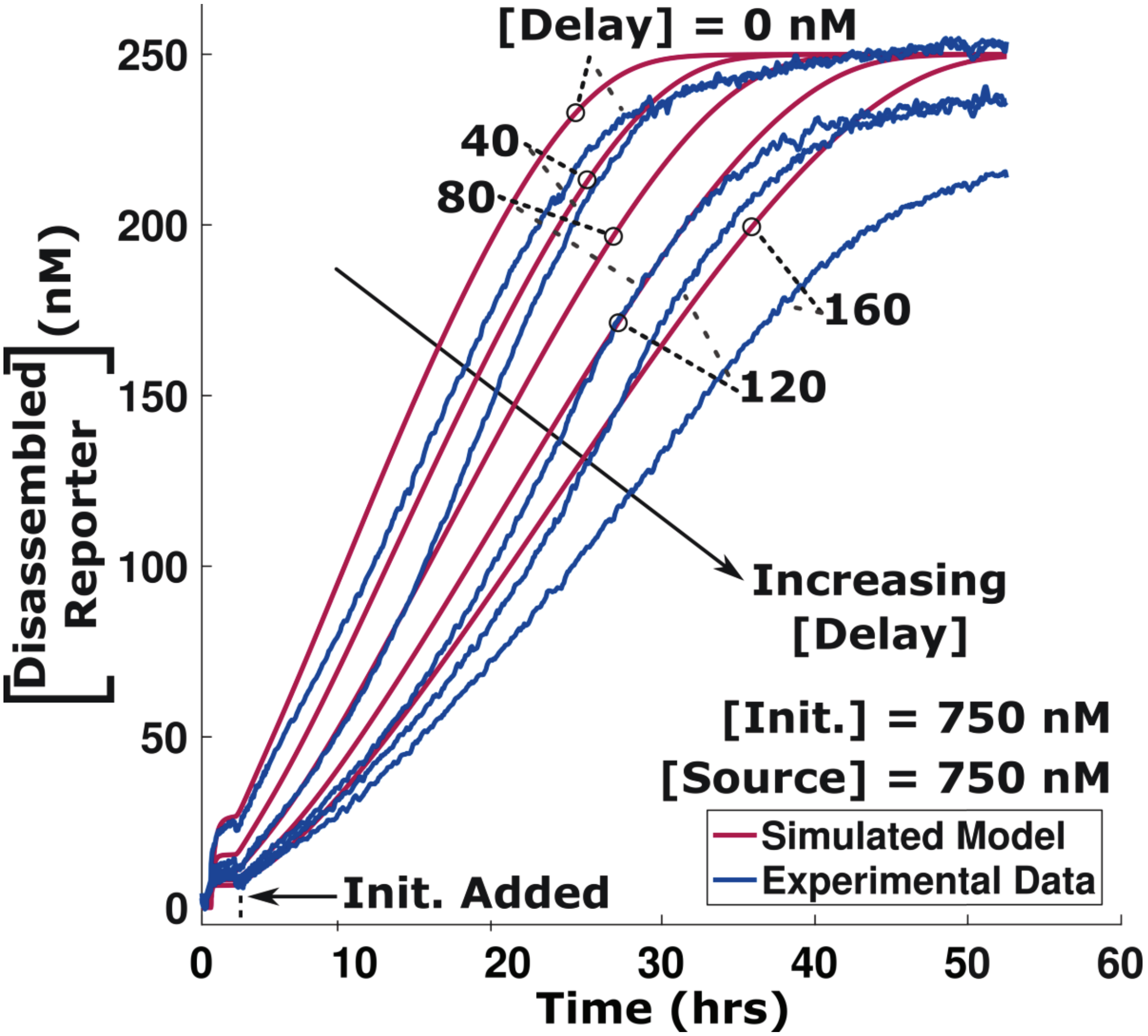


Figure S4.20: Release of the timer circuit Output (as measured by its interaction with the Reporter complex) when Initiator in present. Concentrations of reaction components are listed in Table S4.9. Experiments and simulations show similar trends. As noted above and in Figure S4.18, delayed release is not observed due to Reporter complexes competing with Delay complexes for released Output (Figure S4.17).

Table S4.10: List of sequences and their names used in the experiments.

Strand Name	Role	Sequence
5' TOEHOLD REPORTER		
Rv(W5_)q	5'Rep1 IowaBlackFQ	/5IABkFQ/CA CCACCAAACTT CA
Rb(W5_)f	5'Rep1 FAM	TG AGA TG AAGTTTGGTGG TG/36-FAM/
W5_	5'Output1, No Hairpin	CA CCACCAAACTT CA TCT CA
W5_6.extHP	5'Output1, With Hairpin	CA CCACCAAACTT CA TCT CA TAACACAATCA CA CATCC TTTT GGATG
3' TOEHOLD REPORTER		
Rb5f	3'Rep1 IowaBlackFQ	/56-FAM/TG AAGTTTGGTGG TG AGA TG
Ro5Q	3'Rep1 FAM	CA CCACCAAACTT CA/3IABkFQ/
W_5	3'Output1, No Hairpin	CA TCT CA CCACCAAACTT CA
DECOY DNA		
W3prime_	Decoy DNA	AT AGATTTTAGGG AT CTC AT
W3_	Decoy DNA	AT CCCTAAAATCT AT CTC AT
PolyT20	Decoy DNA	TTTTTTTTTTTTTTTTTTTT
0 bp SOURCE PURIFICATION ADDITIVE		
W5(3)_6	S1.2 No Toehold	CTT CA TCT CA TAACACAATCA CA
RELEASE REACTION STRANDS		
W5_6.extHP	Source1.2 Top/Output	CA CCACCAAACTT CA TCT CA TAACACAATCA CA CATCC TTTT GGATG
Gb5(3)_6_5bp	5 bp Toehold Source1.2 Bottom	TG AGA TG TGATTGTGTTA TG AGA TG AAG
Gb5(3)_6_2bp	2 bp Toehold Source1.2 Bottom	GA TG TGATTGTGTTA TG AGA TG AAG
Gb5(3)_6_0bp	0 bp Toehold Source1.2 Bottom	TG TGATTGTGTTA TG AGA TG AAG
W5(3)_6_.extHP	Initiator1.2 5/2 bp Toeholds	CTT CA TCT CA TAACACAATCA CA TCT CA CATCC TTTT GGATG
W5(3)_6.extHP	Initiator1.2 0 bp Toehold	CTT CA TCT CA TAACACAATCA CA CATCC TTTT GGATG
2-LAYER CASCADE, SECOND LAYER		
W5(3)_6_7.extHP	Source1.2.3 Top	CTT CA TCT CA TAACACAATCA CA TCT CA ACATATCAATT CA CATCC TTTT GGATG
Gb6(3)_7_5bp	5 bp Toehold Source2.3 Bottom	TG AGA TG AATTGATATGT TG AGA TG TGA
Gb6(3)_7_2bp	2 bp Toehold Source2.3 Bottom	GA TG AATTGATATGT TG AGA TG TGA
W6(3)_7_.extHP	Initiator2.3	TCA CA TCT CA ACATATCAATT CA TCT CA CATCC TTTT GGATG
TIMER CIRCUIT		
W5(3)_6.extHP	Initiator1.2 0 bp Toehold	CTT CA TCT CA TAACACAATCA CA CATCC TTTT GGATG
W5_6.extHP	Source1.2 Top/Output	CA CCACCAAACTT CA TCT CA TAACACAATCA CA CATCC TTTT GGATG
Gb5(3)_6_0bp	0 bp Toehold Source1.2 Bottom	TG TGATTGTGTTA TG AGA TG AAG
Tv5	Delay1 Top	CT CA CCACCAAACTT CA
Tb5	Delay1 Bottom	TA TG AGA TG AAGTTTGGTGG TG AG

4.6 Acknowledgements

The authors would like to thank Dominic Scalise for insightful conversations about the work. This work was supported by DOE Early Career Award 221874 (Biomolecular Materials) and NSF-SHF-1527377 awards to R.S. The DOE Early Career Award provided some salary to J.F. and supported the efforts for reaction kinetics modeling and experiments performed specifically to improve models.

CHAPTER 5: MODULAR DNA STRAND-DISPLACEMENT CONTROLLERS THAT DIRECT MATERIAL EXPANSION

Summary

Stimuli-responsive soft materials have shown important applications in the fields of drug delivery, soft robotics, and diagnostics. Chemo-mechanical mechanisms within these materials sense and react to the environment to cause changes to the material's structure that gives rise to its desired function. One particular class of stimuli-responsive materials, DNA-integrated hydrogels, are uniquely adaptable and have shown promise in applications such as molecular sensors or therapeutic delivery devices. We designed novel hydrogel crosslinks built using DNA that change configuration upon reaction with DNA inputs. Crosslinks in an active configuration are able to undergo further reactions that results in high-degrees of hydrogel expansion. DNA-based controllers with the capability to recognize and transduce arbitrary chemical information were implemented to direct changes in the crosslink architecture of the hydrogel, enabling autonomous swelling and shape-change. Since the controllers are realized using DNA strand-displacement circuits, they can perform a multitude of functions from specific molecular recognition to computational logic, where hydrogel swelling is dictated by the presence or absence of specific biomolecular stimuli. Additionally, since DNA strand-displacement circuits are primarily sequence specific, multi-layer and multi-domain hydrogels could be produced where different regions of the same hydrogel are sensitive to different input identities, creating a complex hydrogel network or device with increasing computational and actuatable power. These results serve as a foundation for "smart" metamorphic soft-material devices capable of transducing complex chemical inputs into mechanical actuation.

5.1 Introduction

In biological systems, tissues change shape and behavior in different ways in response to different biomolecular stimuli. In particular, signaling molecules such as growth factors or small molecules are used to exchange information between cells to direct transitions in cell behavior.[220] Signal transduction pathways can amplify and integrate information about these signals and direct a response. This architecture enables a small amount of input signal to trigger responses that may require large numbers of molecules to be activated or transformed, such as in cell migration, differentiation, and growth.[86,221,222] Additionally, cellular systems utilize different biomolecules for different purposes.[223] A protein used for extracellular sensing is not the same protein that activates changes to cellular behavior on the gene-level. This overall architecture shows how complex soft material responses can be programmed with autonomous chemical processes. Developing a similar architecture may make it possible to create a new generation of materials with complex, programmable responses to a variety of inputs.

A fundamental problem in the design of soft materials is the mechanism used to make the material stimuli-responsive. One way to design such responsiveness is to use the stimulus to directly interact with the material to change the material's properties. Such mechanisms have been used to design materials that are responsive to temperature,[224,225] pH,[226] or fluid flow.[227] A growing class of hydrogels change material properties in the presence of biomolecules such as enzymes[228–230] or antibodies[231] by integrating molecules directly responsive to the stimuli into the material. Recently, chemical reaction networks have been used to direct global changes within soft materials such as gelation,[230,232–234] or chemical release.[233,235] One challenge in using molecular circuits to control materials is that the outputs of a chemical circuit must directly control potentially large-scale changes within the material through chemical reactions. Usually, high concentrations of biomolecules have been necessary to produce large-scale changes in materials, which typically cannot be released by circuits.

Chemical amplification processes have been used to amplify signals to direct large-scale changes within materials such as gel-sol transitions[234,236] or signal propagation across long distances.[237] Here we show that we can direct and precisely program millimeter-scale size changes within materials using programmable biomolecular reaction networks that can interpret different input signals, amplify the signals, and direct a size-change response. A chemical amplifier produces a signal of sufficient strength to induce a desired response, which utilizes an existing supply of chemical fuel (Figure 5.1a). Together, the system we develop is a programmable soft robot, where both the robot and the controller, *i.e.*, sensing, signal processing, and mechanical actuation, are fully implemented in biochemistry.

5.2 Materials and Methods

5.2.1 Chemicals and DNA

Acrylamide (Bio-Rad, Cat. No. 161-0100) was solubilized using MilliQ purified water. Rhodamine B-conjugated acrydite monomer was obtained from PolySciences, Inc (Cat. No. 25404-100) and used for fluorescent visualization of the hydrogel. Hydrogels were polymerized using the photoactive initiator Irgacure 2100 (BASF). ATP was purchased from Sigma (Cat. No. A6419) and solubilized to 53 mM using MilliQ purified water. Unmodified and acrydite-modified DNA strands were purchased with standard desalting purification from Integrated DNA Technologies, Inc. Fluorophore- and quencher-modified DNA was purchased with HPLC purification. All DNA was solubilized using 1x TAE buffer (Life Technologies, Cat. No. 24710-030) supplemented with 12.5 mM magnesium acetate tetrahydrate (Sigma, Cat. No. M5661). As described in Figures S5.3, S5.13, S5.17, and S5.19, DNA sequences were designed using NUPACK[116] or adapted from previous literature.[87,107,238,239] Sequences used in this study are found in Table S5.1.

5.2.2 Preparation of DNA complexes

DNA complexes were annealed in 1x TAE buffer supplemented with 12.5 mM magnesium acetate (TAEM) from 90 to 20 °C using an Eppendorf PCR at 1 °C/minute. Hydrogel crosslinker complexes were annealed at a stock concentration of 3 mM per strand while all other complexes were annealed at 100 μM. Hairpin-forming strands were flash cooled on ice for 3 minutes after heating to 95 °C for 10 minutes at a concentration of 80 μM or 400 μM. Hairpin and crosslinker complexes were not further purified. All other multi-strand circuit components (*e.g.*, Source complexes) were PAGE purified after annealing using 15% polyacrylamide gels at 150 V for 3 – 4.5 hours. Immediately prior to PAGE purification, all complexes, with the exception of the ATP sensor complex, were incubated ~16-20 hours with 50 μM of their respective input strand with the toehold removed.[107] The band corresponding to the desired product was excised from the gel and the DNA complex was eluted using TAEM buffer. Fluorophore-/quencher-modified DNA complexes (Reporters) were not PAGE purified after annealing at 50 μM.

5.2.3 Synthesis of poly(DNA-*co*-acrylamide) hydrogel particles

DNA crosslinks were mixed to a final concentration of 1.154 mM with water, 10x TAEM, acrylamide, rhodamine methacrylate, and Irgacure 2100 (75% v/v in butanol). The final concentrations of acrylamide, rhodamine methacrylate, and Irgacure 2100 were 1.41 M, 2.74 mM, and 3% (v/v), respectively. After mixing, the pre-polymer solutions were put under vacuum for 5 minutes. Pre-polymer droplets were prepared using a water-in-oil method (Figure 5.2a). Mineral oil USP (CVS Pharmacy) "wells" were prepared on a cratered parafilm surface and pre-polymer droplets were added using a pipette set to 0.25 μL. Droplets were exposed to 365 nm light using a Benchtop 3UV Transilluminator (UVP) for 1 minute (~4 mW/cm^2) to polymerize and crosslink the particles. Particles were purified from the oil using centrifugation into 1x TAEM and were stored at 4 °C until use, usually within 1 week.

5.2.4 Swelling of DNA-crosslinked hydrogels

Swelling experiments were conducted in 96-well plates (Fisher Scientific) with one particle per well and recorded on a IX73 Olympus fluorescence microscope using a filter specific to rhodamine. The final volume of liquid added to each well varied between 100 – 120 μL, depending on the experiment. Experiments with locked particles were incubated with DNA hairpins (20 μM/hairpin type, 10% terminator) for ~24 hours prior to the addition of Catalyst/Helper strands or circuit complexes. For all experiments with DNA circuits, the Helper strand concentration was 10 μM. Images of each particle were captured every 30 minutes.

5.2.5 Particle area measurement and analysis

Images of the fluorescent particles were considered to be accurate 2D projections of the particle size near the center *xy*-plane. To decrease the sensitivity and bias involved in measuring the diameter, especially of an irregular or non-circular projection, the area of the 2D projection was chosen as the representative variable of particle size and calculated as a function of time for each particle. The area was determined using standard intensity-based thresholding and mask image analysis using MATLAB (Section 5.5.1). Area measurements for each particle were normalized to the initial time point. The curves shown are the average of multiple particles and the data was smoothed using MATLAB's *smooth* function with a smoothing factor of 3.

5.2.6 Fluorophore-quencher assay of DNA strand-displacement controller circuits

An Agilent Stratagene Mx3000 or Mx3005 was used to test the operation of the DNA-based circuits in the absence of hydrogel particles. A reporter complex, using FAM and

IowaBlackFQ fluorophore-/quencher-modified DNA, was designed to have an increase in measured fluorescence upon reaction with DNA strands containing the Catalyst sequence and toehold (Figure S5.14). The measured fluorescence increase was converted into the concentration of Catalyst strand using a calibration curve. The DNA strand-displacement logic circuit was run with 200 nM Source complexes and 200 nM Reporter. Aptasensor circuits were run at 100 nM Source complexes, 100 nM Cofactor strand, and 200 nM Reporter. PolyT$_{20}$ (1 μM) was added to inhibit adsorption to well walls.

5.3 Results

5.3.1 DNA-crosslinked hydrogels as a state-switchable device

To build a material that enables modular control over soft material size change in response to a variety of chemical input stimuli or stimulus combinations, we investigated how to control the size change of polyacrylamide hydrogels crosslinked by hybridized short DNA strands (Figure 5.1b). DNA-crosslinked hydrogels have been designed to respond to an array of inputs including temperature,[240,241] ions,[241,242] nucleic acids,[144] and small molecules[153,243–245] by either de-hybridizing the crosslinks or changing the crosslink's stiffness to direct material swelling or mechanical property changes.

Recently, volumetric expansion of 100-fold or more was achieved through a process by which DNA hairpins are incorporated into the crosslinks at an "active site" *via* insertion polymerization.(Figure 5.1c).[239] The degree of expansion can be controlled by the relative abundance and concentration of hairpins that can incorporate into a crosslink and propagate a site for continued incorporation (polymerizing hairpins) *vs.* hairpins that can incorporate into a crosslink but do not leave a site for continued crosslink extension (terminating hairpins).

The ability to direct extensive mechanical change using DNA hybridization suggests that the swelling could be directed by upstream DNA circuits, such as amplifiers, translators, or logic circuits that would enable expansion in response to different types or concentrations of chemical input stimuli. We hypothesized the development of such a system by modifying hydrogels so that they could be either in an active state, where DNA hairpins could direct hydrogel expansion, or an inactive state, where the conformation of the crosslinks prevents the hairpin polymerization reaction that changes crosslink size (Figure 5.1b, d). Upstream circuits would then control the conformational change of crosslinks that would activate them, similar to biology where transcription factors activate genes for transcription.

In the system we constructed, the switching of the hydrogel from a locked state to an unlocked state (from an inactive to an active state) occurs using a DNA strand-displacement reaction, where a single-stranded domain, a "toehold," is used to initiate and control the kinetics of the strand-displacement process (Figure 5.1d).[96,98] Because the signal that activates the hydrogel is a simple single-stranded DNA molecule, complex input transducing elements such as logic[87,107] and sensing[180,246,247] can be implemented *via* DNA strand-displacement circuits.

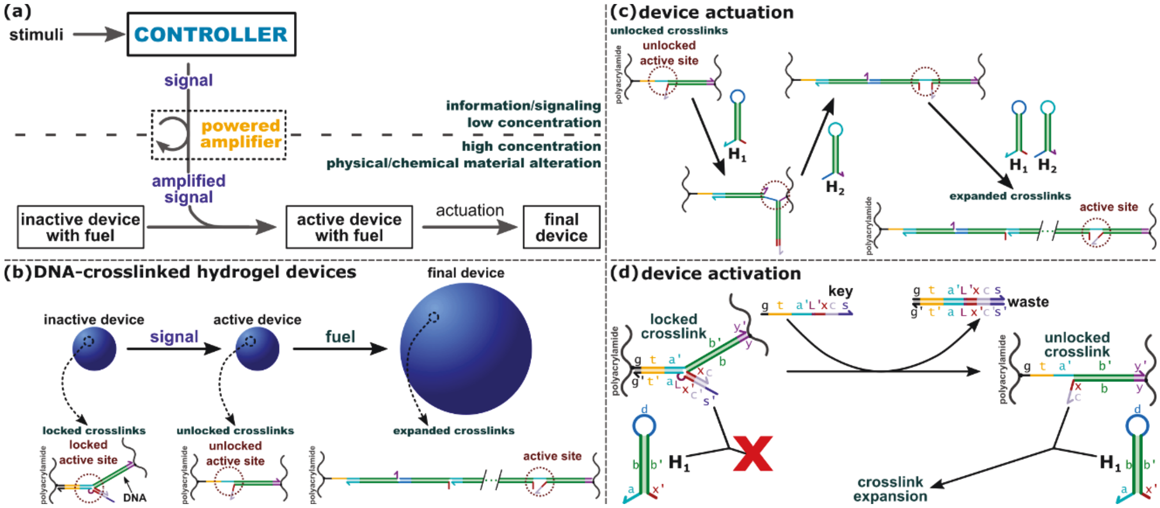


Figure 5.1: Scheme for controlling the expansion state of DNA-crosslinked hydrogels by switching the activity state of hydrogel crosslinks. (a) Control over material changes requires switching from an inactive to an active state. Only from the active state can fuel be utilized to actuate the device. Amplification of environmental stimuli direct the state switching. (b) Polyacrylamide hydrogels with DNA crosslinks can be switched from a locked state to an unlocked state *via* a strand-displacement reaction. (c) Once unlocked, DNA hairpins can be incorporated into the crosslinks to drive hydrogel swelling. (d) Device activation is defined by the unlocking of the crosslinks. Key strands remove the lock on the crosslink through a strand-displacement reaction. DNA domains are drawn to scale.

5.3.2 DNA-crosslinked hydrogel particles as a model swelling system

In order to understand how the kinetics of DNA circuits control the kinetics of hydrogel size change, we sought to monitor the kinetics of both the circuit and swelling processes. One challenge in characterizing these changes is that different portions of a hydrogel can swell differentially depending on its geometry. Corners, for example, have greater interfacial exposure to the surrounding solution. Anisotropies in swelling can also be magnified because changes in size can influence the interfacial surface and hydrogel pore size, which could improve DNA hairpin transport into and through the hydrogel. To minimize these effects, we chose to characterize size change in spherical hydrogel particles, which are rotationally symmetric. DNA-crosslinked polyacrylamide hydrogel particles were synthesized by photopolymerizing pre-polymer droplets suspended in mineral oil (see Section 5.2 and Figure 5.2a). Rhodamine-B methacrylate was added to the pre-polymer solution so that the polymerized particles could be visualized using fluorescence microscopy. To measure particle size, we use the area of the 2D projections of the particles in the fluorescence micrographs (Section 5.5.1, Figure S5.1). The average particle radius after synthesis was measured to be 570 ± 50 μm (Section 5.5.2). The projection of 93.4% of the particles was measured to be at least 90% circular in shape, and circularity was maintained after hydrogel swelling (Figure S5.2).

First, we verified that hydrogel particles without crosslink locks swell when mixed with DNA hairpins. Particles were incubated in a solution containing 20 μM of H_1 and H_2 hairpins with 10% terminator hairpins. The area of the hydrogel projection increased linearly until reaching a stable final size (Figure 5.2b). The intensity of the particles became non-uniform during the swelling process, indicating that complete swelling was dependent upon the diffusion of hairpins to the interior of the hydrogel particles (Figure 5.2b). The total intensity of a swelling hydrogel decreased with increasing particle size due to the decreasing density of Rhodamine (Figure S5.1). We chose 20 μM hairpins and 10% terminator due to the relatively fast swelling to a final, stable size coupled with a high-magnitude change in area of ~300% (Figure 5.2b, Figures S5.3 – S5.5).

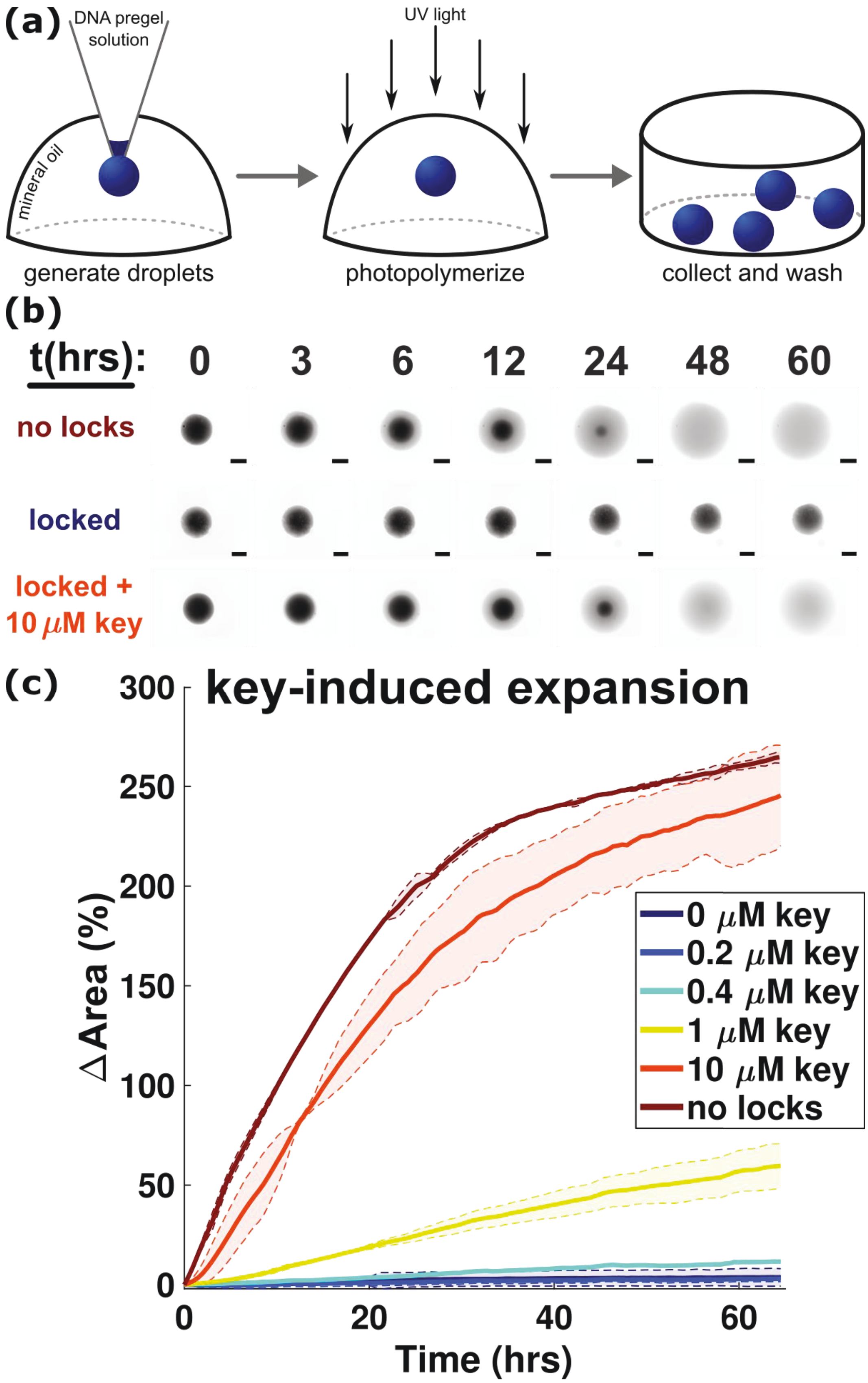

(a)
DNA pregel solution
mineral oil
generate droplets
UV light
photopolymerize
collect and wash
(b)
t(hrs):
0
3
6
12
24
48
60
no locks
locked
locked +
10 μM key
(c)
key-induced expansion
ΔArea (%)
0
50
100
150
200
250
300
Time (hrs)
0
20
40
60
0 μM key
0.2 μM key
0.4 μM key
1 μM key
10 μM key
no locks

Figure 5.2: Hydrogel particle generation and unlocking *via* key strands. (a) Hydrogel particles are prepared by pipetting droplets of pre-polymer acrylamide-DNA crosslinker solution into mineral oil and polymerizing with UV light. Rhodamine-B is incorporated into the acrylamide backbone for fluorescence imaging (see Section 5.2). (b) Fluorescence micrographs of particle area used to track swelling. Particles without locks swell when hairpin fuel is added, but particles with locked crosslinks show less than 4% change in area over 70 hours (Figure S5.6). Addition of key strands to locked particles with hairpins initiates swelling. Images are scaled to initial particle size. Scale bars: 500 μm. (c) The change in area of the 2D projection of each particle over time tracking swelling as a function of the concentration of the key strand. Hairpin concentration is 20 μM per type with 10% terminator. Curves are the average of 2 particles; shaded regions show 95% confidence intervals as determined by standard deviation.

5.3.3 Particles activated using high concentrations of key strand

Next, hydrogel particles were prepared with locked crosslinks. When these particles were incubated with hairpin fuel at 20 μM and 10% terminator, the degree of swelling was drastically reduced. Only 3.7±4.5% change in area was observed compared to 260±2% for crosslinks without locks over 60 hours (Figure 5.2b, Figure S5.6). Increasing the hairpin concentration to 200 μM with 10% terminator hairpins showed only 3.4±0.6% swelling (Figure S5.7). Thus, hairpin fuel cannot trigger significant expansion of a hydrogel with locked crosslinks.

We next tested whether adding key strand could direct the hydrogels to expand in the presence of hairpin fuel. By removing one of the strands of the locked crosslink, the now unlocked crosslink and expansion hairpins can interact (Figure 5.1d). We first incubated the locked particles with the hairpin fuel for 24 hours, then added different concentrations of the key strand to different particles. After a 2 – 5 hour period in which little expansion was observed, the area of hydrogel particles increased at a roughly linear rate. The swelling rate slowed after a period of time that depended on the concentration of key strand (Figure 5.2d). At 200 nM key strand, little to no swelling was observed. One reason for this behavior could be that the number of crosslinks unlocked by key strands, compared to crosslinks that are still locked, are limiting both swelling speed and the final particle size. To test this possibility, we prepared hydrogel particles with varying proportions of two different types of crosslinks with two different sequences. One crosslink is capable of being expanded *via* added DNA hairpins while the other crosslink does not have matching sequences for expansion (Figure S5.8). Both the kinetics and, primarily, the final size was determined by the proportion of crosslinks that were expandable.

5.3.4 Catalytically unlocking crosslinks decreases required input concentration

Key strands can enable an inactive hydrogel to become active and can, in the presence of the hairpin fuel, act as a stimulus that directs size change. However, a relatively high concentration of key strand, on order 10 μM, is required. We next sought to design materials that could respond to lower concentrations of a signal by using a molecular amplifier that would activate hydrogels in response to a much smaller concentration of input molecules. Such an amplification process would make it possible to use molecular circuits such as DNA strand-displacement reactions, which have been demonstrated for input concentrations of 1 nM to 1 μM,[87,101,107–111,181,187,191,246,248,249] to process information about the environment and release a DNA strand output to initiate the amplification process. Such signaling processes could make it possible to engineer large-scale size change in response to smaller concentrations of inputs, logical combinations of nucleic acid signals, or other molecules using sensing molecules such as aptamers.[107,247] Such an architecture would also mean that only small amounts of the molecules in the signaling cascade would need to be present, making it much more practical to implement complex processing systems requiring many different species.[87,191]

To develop such a system, we used a catalytic strand-displacement circuit[91,96,109,111,181,191,249] to amplify an input signal and produce a high degree of hydrogel activation similar to when a high concentration of key strand is used. In catalytic DNA strand-displacement circuits, a small concentration of input initiates a cascade in which the input interacts with a complex to release an output, and then the input is re-released to participate in subsequent output releasing processes. High turnover rates (100-1000) of the input can result in signal amplification.[91,96,109,111]

We built a catalytic process for unlocking crosslinks in response to a small concentration of input by replacing the key strand with Catalyst and Helper strands (Figure 5.3a). The Catalyst strand can bind to a crosslink, partially releasing the lock. The Helper strand can then bind to the Catalyst-crosslink complex, releasing both the unlocked crosslink and the Catalyst while producing a waste complex. A small concentration of Catalyst, which can be reused in multiple unlocking reactions, along with a larger concentration of Helper, which is consumed in an unlocking process, should together unlock crosslinks.

Without the Helper strand, a Catalyst is able to react with and unlock a crosslink, but will not be released to interact with further crosslinks (Figure 5.3a). Therefore, the amount of swelling observed in response to a particular concentration of the Catalyst strand alone should be roughly the same as the amount of swelling observed in response to that same concentration of key strand. When locked hydrogel particles were incubated with Catalyst and the hairpin fuel, we found that 10 μM of Catalyst was needed to achieve the high-degree change in area of 270±1% over 60 hours, and less than 2% was observed at Catalyst concentrations of 100 nM or below (Figure S5.9).

Similarly, when the particles are incubated with Helper and the hairpin fuel alone, limited to no unlocking and therefore little or no expansion should occur because there is no toehold on the lock where the Helper strand can bind to initiate the unlocking process. As expected, locked hydrogel particles incubated with Helper strands and hairpin fuel were found to have less than 8±2% change in area when the [Helper] was at or below 1 μM (Figure S5.10). At 10 μM, the change in particle size rose to 47±4% after 40 hours, still well below the 200 – 250% increase in area observed with the key or Catalyst strands alone at 10 μM (Figure 5.2d, Figure S5.9).

In contrast to the limited expansion observed with Helper strands and hairpin fuel alone, expansion at rates similar to the unlocked particles was observed when as little as 100 nM Catalyst was added to the Helper and hairpin fuel (Figure 5.3c). The use of more Catalyst strand does not significantly speed up the swelling process when 10 μM Helper is present, consistent with the idea that Catalyst strands are each capable of initiating multiple unlocking reactions. In contrast, 1 μM Helper strand in the presence of large amounts of Catalyst was not sufficient to induce significant expansion over the 27 hour period during which 160±11% expansion of particles was induced by 100 nM Catalyst and 10 μM Helper strand (Figure 5.3c, Figure S5.11).

As a secondary method to control the degree of swelling of locked hydrogels using Catalyst and Helper strands, we tested the kinetics of swelling at hairpin concentrations in the range 1 – 200 μM per type (Figure S5.12). Interestingly, the kinetics of swelling was slowed when the hairpin concentration was decreased below or increased above 20 μM. Above 20 μM, this difference in kinetics could be caused by the binding between hairpins and Catalyst/Helper strands due to the sequence complementarity of the active site domains, decreasing effective concentration of Catalyst and Helper, and thus decreasing the number and rate of crosslink activation. Below 20 μM, the lower hairpin concentration decreases the resulting final length of the expanded crosslinks, thus limiting the final size of the swollen hydrogel state.

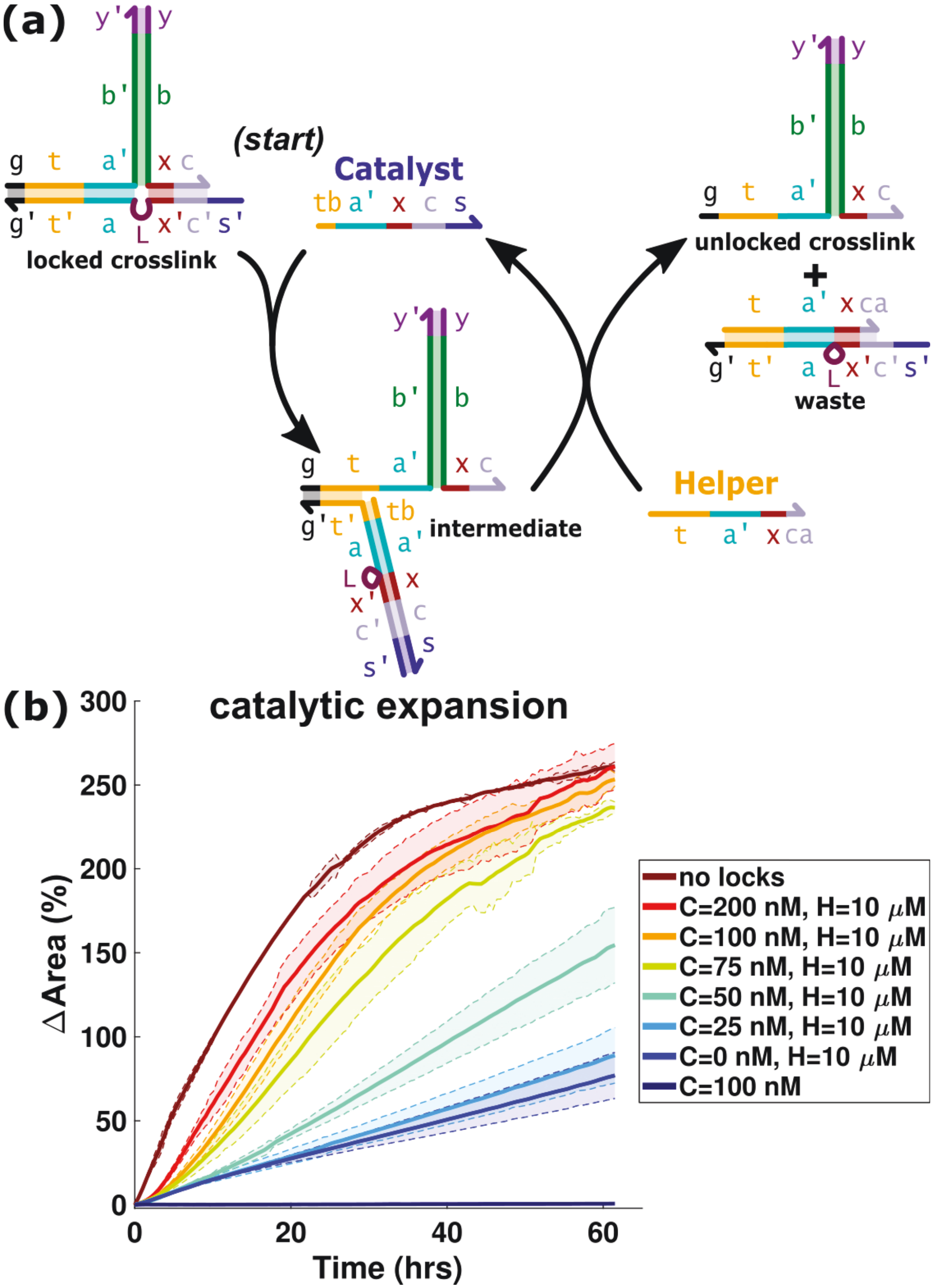

(a)
y'
y
b'
b
g
t
a'
x
c
g'
t'
a
L
x'
c'
s'
locked crosslink
(start)
Catalyst
tb
a'
x
c
s
unlocked crosslink
+
waste
intermediate
Helper
t
a'
x
ca
(b)
catalytic expansion
ΔArea (%)
Time (hrs)
0
50
100
150
200
250
300
0
20
40
60
no locks
C=200 nM, H=10 μM
C=100 nM, H=10 μM
C=75 nM, H=10 μM
C=50 nM, H=10 μM
C=25 nM, H=10 μM
C=0 nM, H=10 μM
C=100 nM

Figure 5.3: Catalytically unlocking crosslinks enables sub-100 nM trigger concentrations. (a) A Catalyst strand unlocks the crosslink *via* toehold-mediated strand-displacement to form an intermediate complex and opens the active site for hairpin incorporation. A Helper strand reacts with this intermediate to release the Catalyst strand, allowing a single Catalyst strand to unlock multiple crosslinks. (b) Different concentrations of Catalyst and 10 μM Helper strands are added to locked particles pre-incubated with DNA hairpins for 24 hours. Helper strands can unlock crosslinks alone, but at a significantly slower rate than in the presence of Catalyst. Catalyst alone (100 nM) does not induce swelling. Curves are the average of 2 – 4 particles with shaded regions representing 95% confidence intervals as determined by standard deviation.

5.3.5 Triggering mechanical change in response to small molecule inputs using aptamer sensors

A catalytic process for unlocking crosslinks makes it possible to trigger expansion using only a small concentration (10 – 100 nM) of a catalyst DNA strand. If a catalyst strand was the output of a DNA strand-displacement circuit, hydrogel expansion could instead be triggered in response to the inputs of the circuit that produces the catalyst output. Strand-displacement circuits can be designed to release a strand only in response to certain concentrations of inputs,[87,107] in response to a Boolean function of multiple inputs,[87,107,108,187] small molecule inputs,[180,246,247] or after programmed time delays.[101]

To demonstrate how strand-displacement circuits can act as controllers for the swelling of a material, we designed different strand-displacement circuits that produced the Catalyst output and coupled them to the expansion process. We first designed an aptasensor circuit that translates the small molecule ATP into a strand containing the Catalyst sequence (Figure 5.4a, Figure S5.13).[246] In the absence of ATP, the Catalyst sequence is partially sequestered in double-stranded form, preventing it from interacting with the hydrogel crosslinks. The sensor is designed to require a Cofactor strand as a secondary input to ensure that the sensor complex maintains its form and stability in the absence of ATP.[246] When ATP and the Cofactor are both present in solution, the Catalyst sequence is exposed and capable of unlocking crosslinks to induce swelling.

To test that the aptasensor operated as designed prior to integration with hydrogels, the concentration of circuit output over time was measured using a fluorophore-quencher reporting assay. In this assay, free strands containing the Catalyst sequence can react with a Reporter complex, leading to an increase in measured fluorescence that is proportional to the concentration of free Catalyst in solution (Figure S5.14). When ATP was added to 100 nM circuit complexes and 100 nM Cofactor, the fluorescence intensity increased in relation to the concentration of added ATP (Figure 5.4b). Catalyst concentrations nearing ~75 nM, the required amount for high degrees of hydrogel particle swelling, was detected only for ATP concentrations above 500 μM, while less than 40 nM Catalyst was produced at 100 μM ATP. When no ATP was present, less than 20 nM Catalyst was produced over 45 hours. This release of Catalyst was likely due to spontaneous dissociation of the aptamer strand from the ATP sensor when the Cofactor was bound or due to undesired "leak" reactions inherent to DNA strand-displacement reactions.[96,99,109,250]

When locked hydrogel particles were incubated with the ATP-driven controller circuit, the amount of particle expansion depended on ATP concentration (Figure 5.4c). Interestingly, the dose-response of ATP to the particles was somewhat digital. ATP concentrations below 500 μM did not significantly increase the swelling rate over a baseline swelling rate observed at no ATP added, whereas swelling rates were similar for ATP concentrations of 500 μM and above. Importantly, the addition of the ATP-sensing circuit to the amplifier circuit, crosslink locks, and hairpin fuel did not change the "leak" swelling rate (Figure S5.15), suggesting that the sensing portion of the controller, the amplifier, and hydrogel expansion are modular units.

The aptamer sequence and the hydrogel crosslinks have no sequence overlap, suggesting that it would be straightforward to replace either the aptamer sensor or the crosslinks to direct hydrogel expansion in response to a variety of chemical inputs, or to direct swelling of different types of DNA-crosslinked hydrogels in response to the same stimuli.

In addition to operating at 100 nM/complex, the circuit was also tested at 200 nM/complex (Figure S5.16), but decreased sensitivity to the presence of ATP was observed at long times due to an increased release of Catalyst in the absence of ATP. This limited the temporal use of the aptasensor circuit with 200 nM/complex, demonstrating the importance of using robust DNA strand-displacement circuit architectures with operating concentrations optimized to decrease the effects of leak reactions.

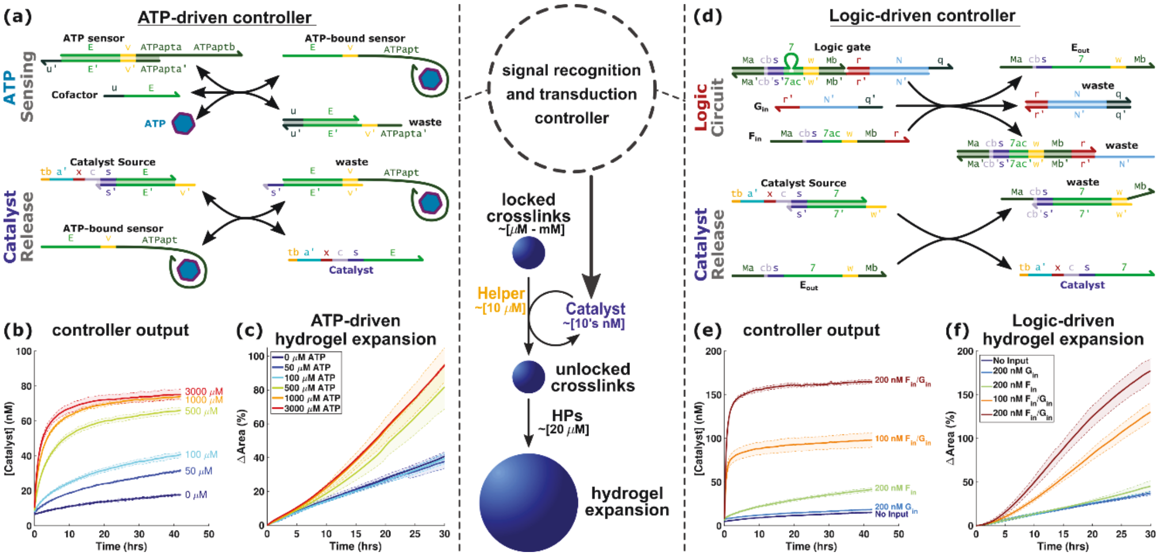


Figure 5.4: A controller for directing hydrogel expansion in response to different types of chemical inputs. (a) An aptamer-based strand-displacement circuit that releases Catalyst strand in response to ATP. (b) Aptamer-based strand-displacement circuit kinetics measured using fluorescence reporting measurements. ATP sensor, Cofactor, and Catalyst Source concentrations were initially 100 nM. (c) ATP controlled hydrogel particle swelling mediated by the ATP-sensitive strand-displacement circuit in (a), catalytic hydrogel unlocking process in Figure 5.3a, and hairpin-polymerization process. After 30 hours, particles swelled 2 – 3x more in the presence of ATP compared to without ATP. Sensor, Cofactor, and Catalyst Source were at 100 nM; Helper strand was at 10 μM. (d) DNA strand-displacement AND logic circuit. Two inputs, F_{in} and G_{in}, are required for Catalyst strand release. (e) Kinetics of the logic circuit as in (d) showing Catalyst is released from the circuit to significant levels only in the presence of both F_{in} and G_{in} as measured by a fluorescent reporter. The logic gate and Catalyst Source were at 200 nM. (f) Locked DNA-crosslinked particle expansion in the presence of different concentrations of F_{in} and G_{in} inputs. The logic gate and Catalyst Source were at 200 nM; Helper strand was at 10 μM. Curves are the average of 3 wells (b, e) or 2 – 6 particles (c, f) and shaded regions represent 95% confidence intervals as determined by standard deviation.

5.3.6 Triggering hydrogel actuation in response to logic combinations of chemical inputs

We next tested whether hydrogel expansion could be directed in response to specific combinations of molecular inputs at a range of input concentrations. It has been previously shown that hydrogels can be engineered to melt only in response to specific combinations of chemical inputs.[232,243,251] However, in these systems, the inputs were required to interact directly with the crosslinks, limiting the kinds of chemical inputs that could be allowed and requiring very high concentrations of inputs to elicit an observable response. We asked whether the use of a modular controller to detect signals would enable control over hydrogel size change rather than melting and whether *in situ* signal amplification would make it possible to direct size change with only very small concentrations of input signals.

To build a modular logic circuit that changes hydrogel size without crosslink dissociation, we modified a previously developed AND logic gate architecture to release Catalyst in response to two different DNA inputs (Figure 5.4d, Figure S5.17).[107] When both inputs, F_{in} and G_{in}, are present, the circuit releases a strand that contains the Catalyst sequence; very little Catalyst is released otherwise. This controller architecture keeps output production at a minimum when no or just one input is added, as demonstrated using the fluorophore-quencher reporting assay to track Catalyst release (Figure 5.4e). When both inputs are present, the concentration of the Catalyst output depends upon the concentrations of both inputs because an amplification step was not used to amplify the output strand signal produced by the logic gate (Figure 5.4d). While an output amplification step is required to achieve digital logic, such circuit additions tend to increase the propensity for leak reactions that would increase the amount of output Catalyst when the circuit is in the OFF state.[87,109]

Locked hydrogel particles incubated with the logic circuit in the absence of logic circuit inputs showed little to no increase in swelling beyond that of 10 μM Helper strand alone (Figure 5.4f, Figure S5.18), demonstrating that like the ATP-sensing controller, this controller operates modularly with respect to the amplifier and unlocking chemistry. When both inputs to the logic controller were present, swelling rates were comparable to those achieved in unlocked hydrogels or with large Catalyst concentrations. Together these results demonstrate how the logic controller allows a set of inputs dictate the degree of hydrogel swelling.

As expected from the reporting assay (Figure 5.4e), some swelling was observed when input F_{in} was the only input to the logic circuit, likely due to some sequence similarity between F_{in} and the Catalyst strand, notably the toehold domain *s* and *cb* (Figure S5.17). To further demonstrate the significant difference in swelling between the ON state and OFF state of the circuit, we devised a logic swelling truth table (Figure 5.5).

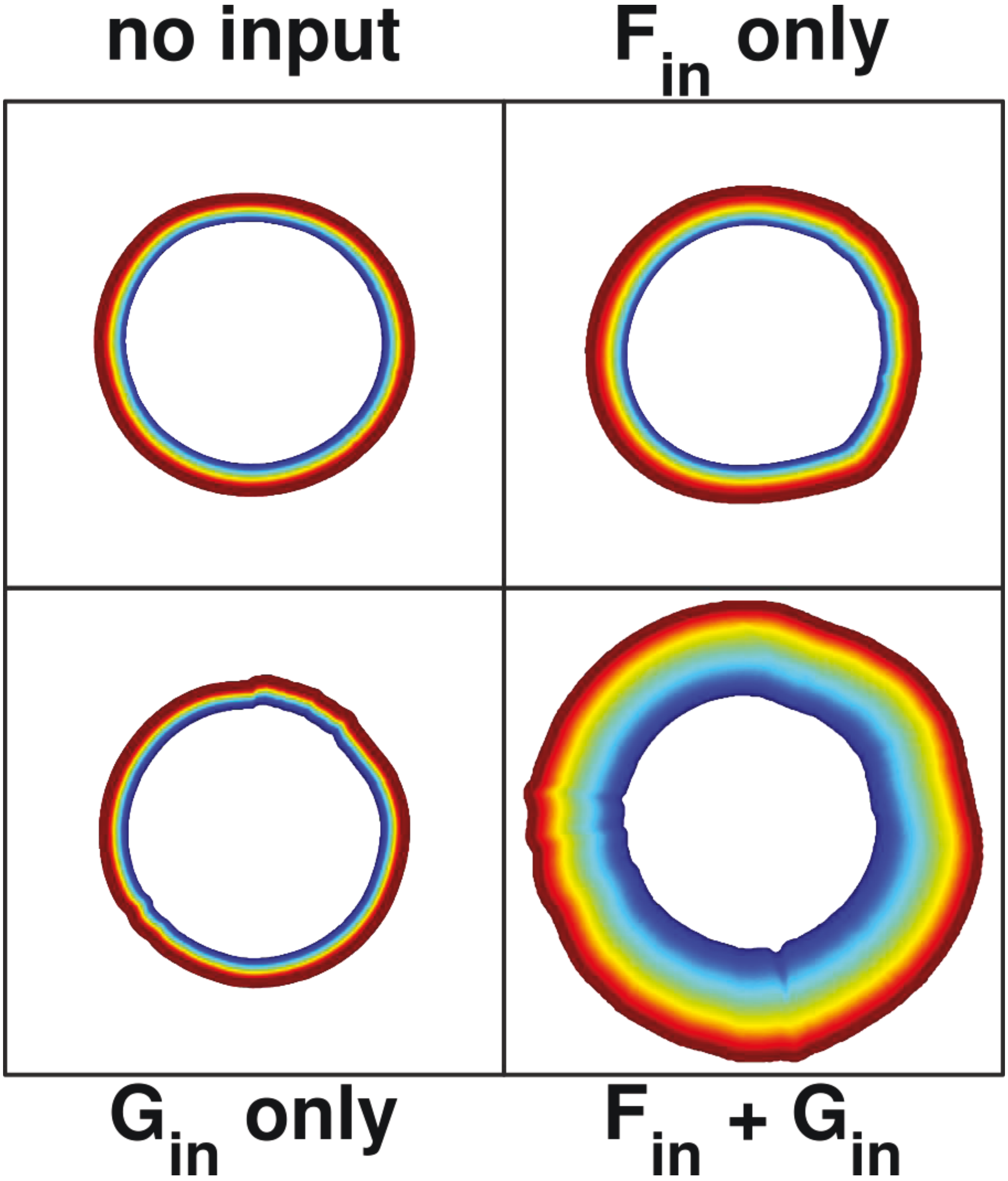


Figure 5.5: Logic-directed swelling of DNA crosslinked-hydrogel particles. Each colored line represents the boundary of each particle in the 2D projection at time points in 30 minute intervals, from blue to red, over 30 hours. Schematics of each particle are scaled such that the initial area of each particle is equal.

5.4 Conclusion

Here we have shown how to use catalytic amplification of a small concentration of a trigger molecule to direct a dramatic change in material size. Further, because this triggering is mediated by a chemical circuit, most of whose components do not interact directly with the material, it is straightforward to create components where different stimuli can induce hydrogel expansion.

The modular design of our system also suggests that there is no immediate barrier to increasing the complexity of the circuits that direct responses. Existing strand-displacement circuits can interpret tens of different signals, amplify signals by 1,000-fold, control timing, or detect single base changes in inputs. Multiplexing circuits to multiple types of crosslinks within multiple domains of a material could also make more complex responses and transformations possible. In combination with sensors, it might also be possible to develop chemo-mechanical feedback processes. Such mechanisms suggest a route to the design of soft materials with a high complexity of response and actuation, paving a path to the design of complex active materials, *i.e.*, robots, where the control system is implemented entirely in chemistry.

5.5 Supplemental Information

5.5.1 Measuring the area of a particle's 2D fluorescence projection

The area of the 2D projection of each particle in the fluorescence micrographs (Figure S5.1) was calculated using custom built MATLAB algorithms. Thresholding was used to determine the boundaries of the particles. The threshold value was calculated using the following method:

1. Normalize the image to the highest and lowest intensity.

$$normImage = \frac{Image - min(Image)}{max(Image) - min(Image)} \tag{S5.1}$$

2. Use MATLAB's built-in global threshold calculator *graythresh*.

$$globalThresh = graythresh(normImage) \tag{S5.2}$$

3. Adjust the global threshold for the non-normalized image and image-specific adjustments.

$$\begin{aligned} initialThresh = globalThresh * (max(Image) - min(Image)) \\ + min(Image) \end{aligned} \tag{S5.4}$$

$$intThreshOrig = initialThresh\ of\ 1st\ image\ in\ time\ series \tag{S5.5}$$

$$\beta = \frac{mean(Image) *\ \alpha}{intThreshOrig} \tag{S5.6}$$

$$1.1 \leq \alpha \leq 3 \tag{S5.7}$$

$$particlePixels = Image >\ \beta * initialThresh \tag{S5.8}$$

where α corresponds to a manual input that was adjusted until a close match between the visible particle boundaries and the calculated boundaries was found. After determining the pixels corresponding to the particle, the particle's area and boundaries were extracted using the functions *regionprops* and *bwboundaries*.

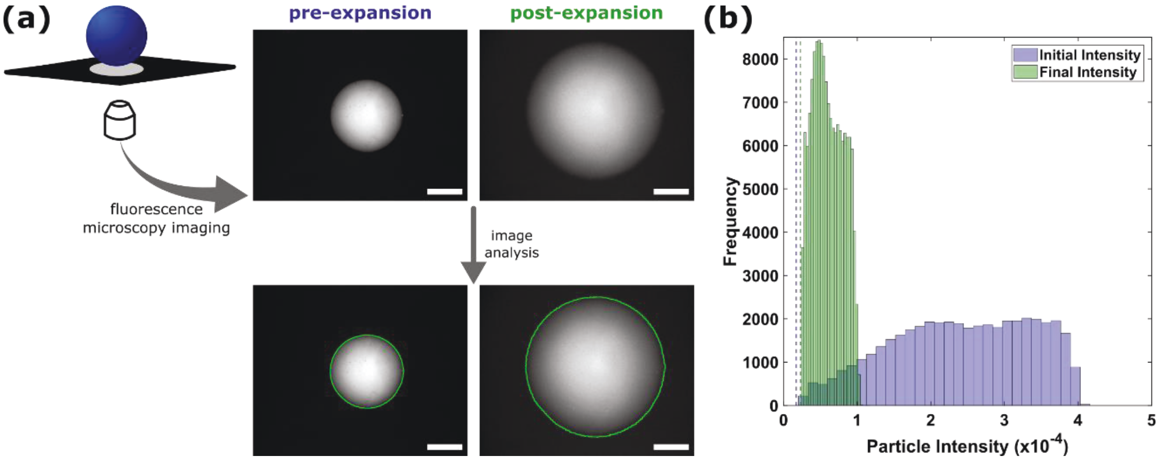


Figure S5.1: Measuring the area and intensity of a particle's 2D fluorescence projection. (a) Particles were imaged using fluorescence microscopy and the area of the particle as seen in the 2D micrograph was determined using MATLAB (Section 5.5.1). The green line around the particle in the bottom images represent the calculated boundaries of the particle. The area of the pixels within this boundary is the calculated particle projection area. Image intensities are scaled based on each image's minimum and maximum intensity. Scale bars are 500 µm. (b) Histogram of the intensity of each particle (pixels within the green boundary) shown in (a). As the particles expand, the intensity of the particle decreases because the density of rhodamine fluorophores decreases.

5.5.2 Calculating particle radii and circularity

The radius of a particle was determined by calculating the average distance from the particle boundaries (Section 5.5.1) and the centroid of the particle. Particle centroids were determined using the function *regionprops* and the distance, in pixels, was calculated using the function *pdist2* in MATLAB. The average radius in pixels was converted to microns by multiplying the radius by the image's pixel size (4.44 micrometers/pixel). The average radius of 196 particles was calculated to be 570±50 μm (Figure S5.2a).

The circularity score of each particle (Figure S5.2b) was calculated using the area of the particle and perimeter of the particle boundary. Both particle area and perimeter were calculated using the *regionprops* function in MATLAB. The circularity score is then calculated using equation 5.9. A score of 1 indicates a perfect circle.

$$Circularity = \frac{4\pi Area}{Perimeter^2} \qquad (S5.9)$$

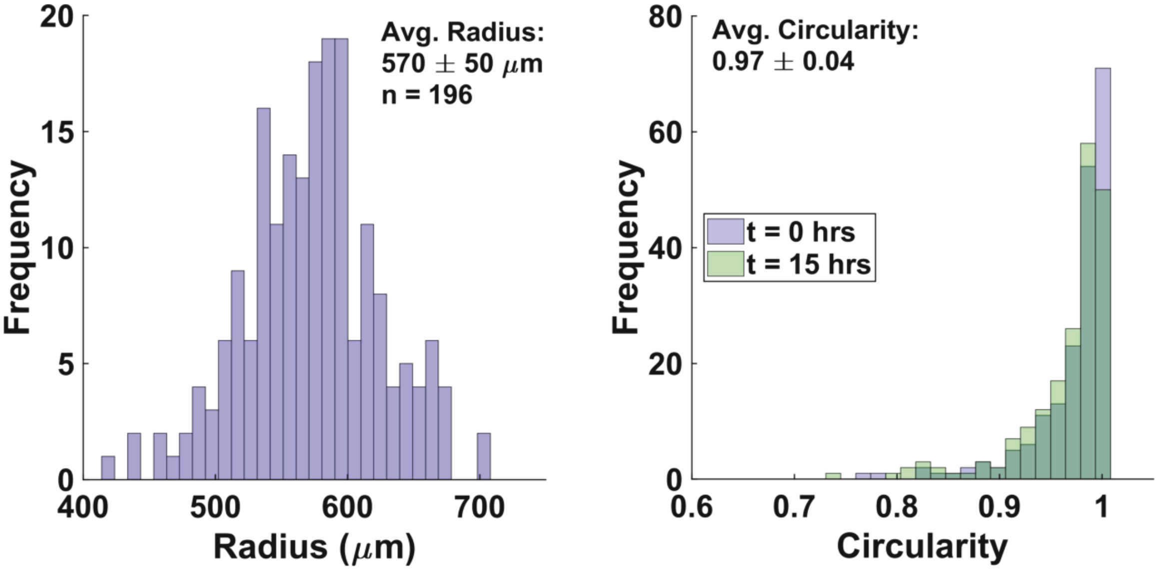


Figure S5.2: Average radii and circularity of DNA-crosslinked hydrogel particles. (a) Distribution of calculated Rhodamine-labeled poly(DNA-*co*-acrylamide) particle radii as described in Section 5.5.2. (b) Distribution of 196 particle's circularity scores as described in Section 5.5.2. Overall, particles maintain their projected 2D circular shape after expansion for 15 hours. A score of 1 defines a perfect circle.

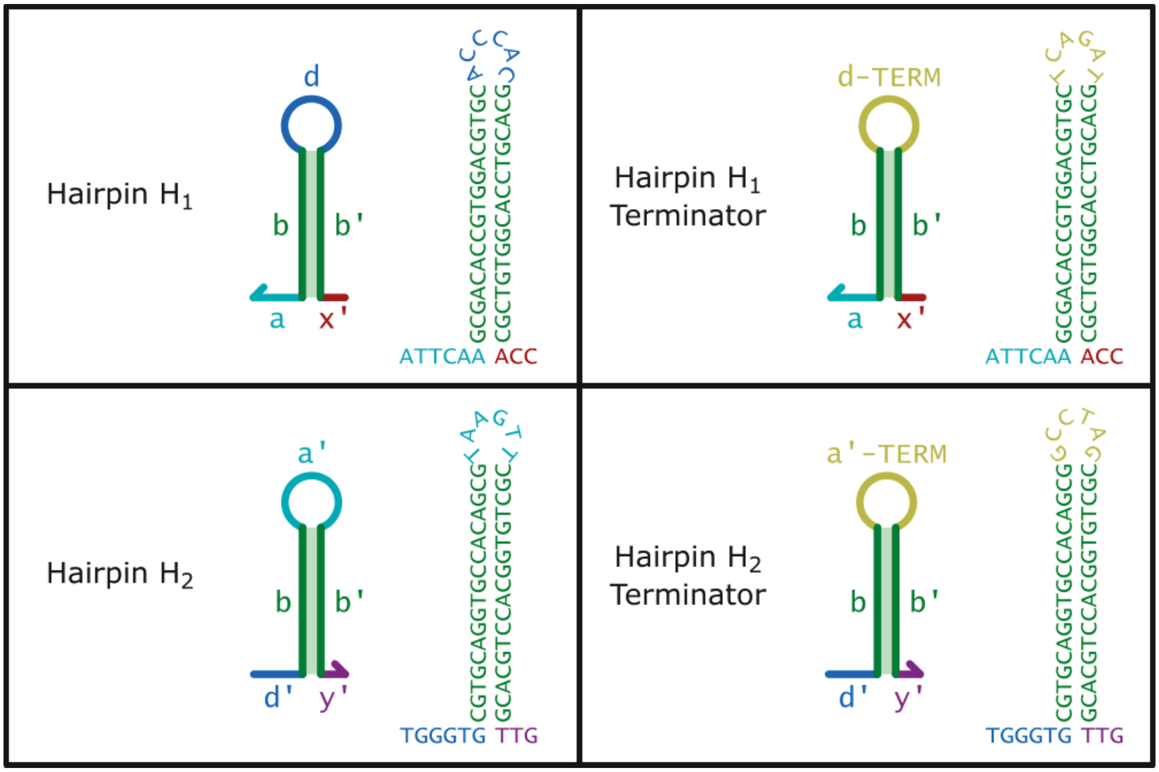


Figure S5.3: DNA hairpin monomer sequences.[239] Terminator hairpins share the same sequence as their non-terminator counterpart except for the loop region. The terminator domains prevent the next hairpin from inserting into the crosslink once the terminator hairpin is incorporated.

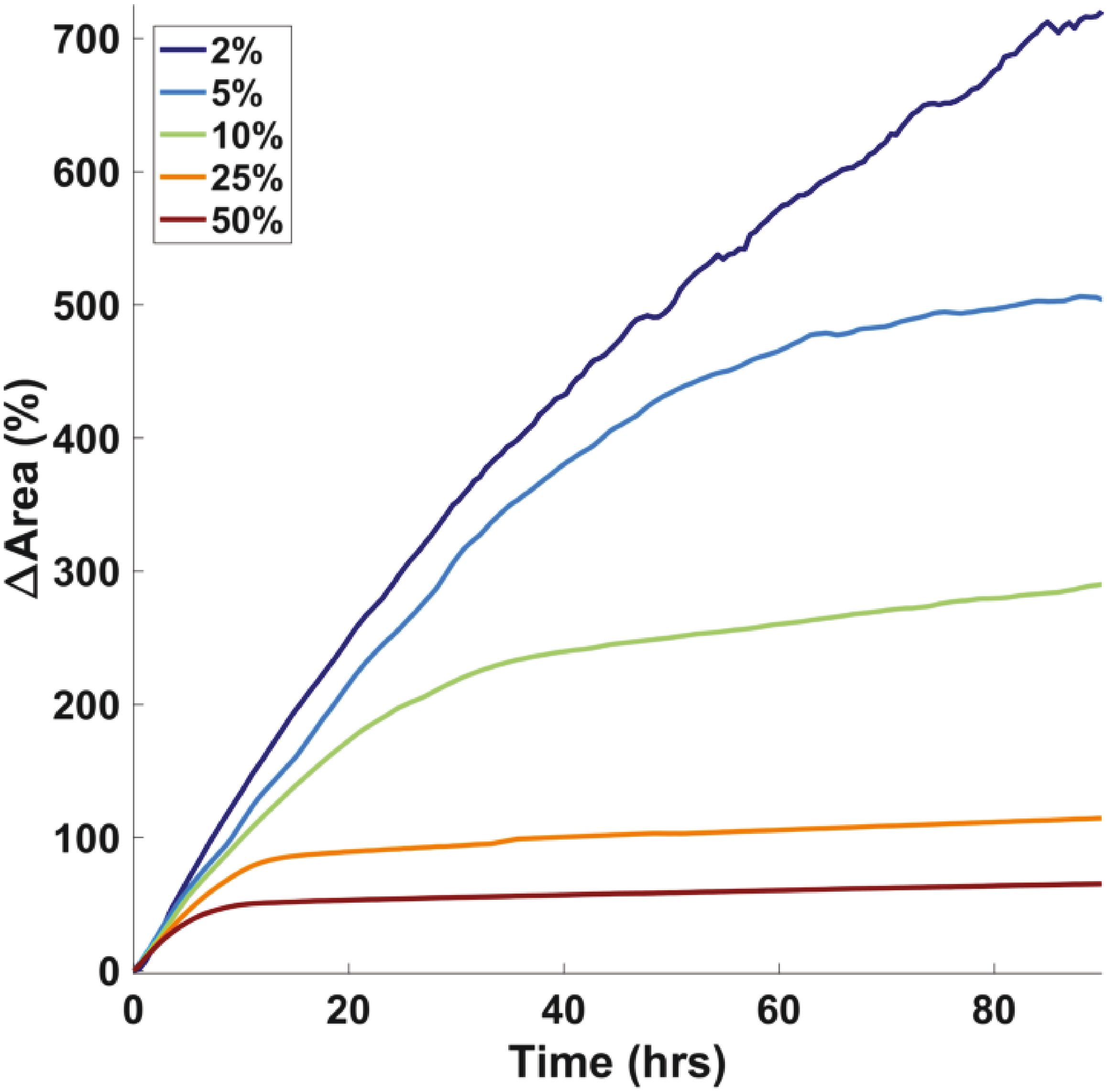


Figure S5.4: Effect of hairpin terminator on swelling of DNA-crosslink hydrogels without locks. Particles without locks were incubated with DNA hairpins at 20 μM total per type at varying fractions of hairpin terminator. The equilibrium size of the particles decreased with increasing hairpin terminator ratio. Ten percent hairpin terminator was chosen for all future experiments due to the relatively quick rise to a stable swelling state coupled with a moderate size increase. Terminator hairpins at 2 and 5% showed increased swelling sizes compared to 10%, with 2% increasing beyond over 700% change in area after 90 hours. Each curve represents 1 particle or the average of 2 particles.

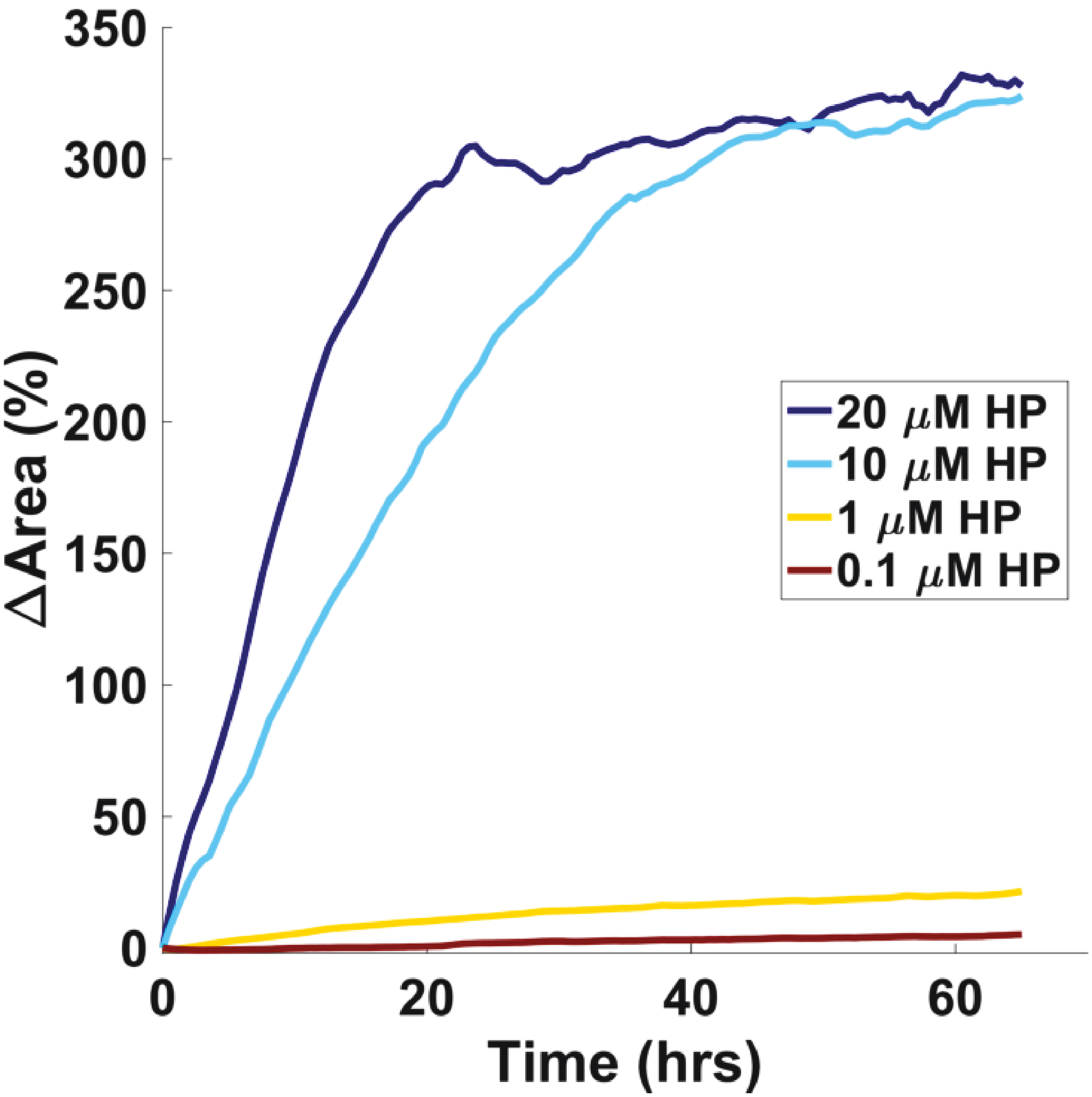


Figure S5.5: Particles prepared without crosslink locks incubated with varying concentrations of hairpin (concentration per type is shown). In all cases, the terminator percentage is fixed at 10%. The kinetics of swelling increased with increasing hairpin concentration. Swelling approached the final swelling state within 24 hours at 20 μM per hairpin type and within 45 hours at 10 μM.

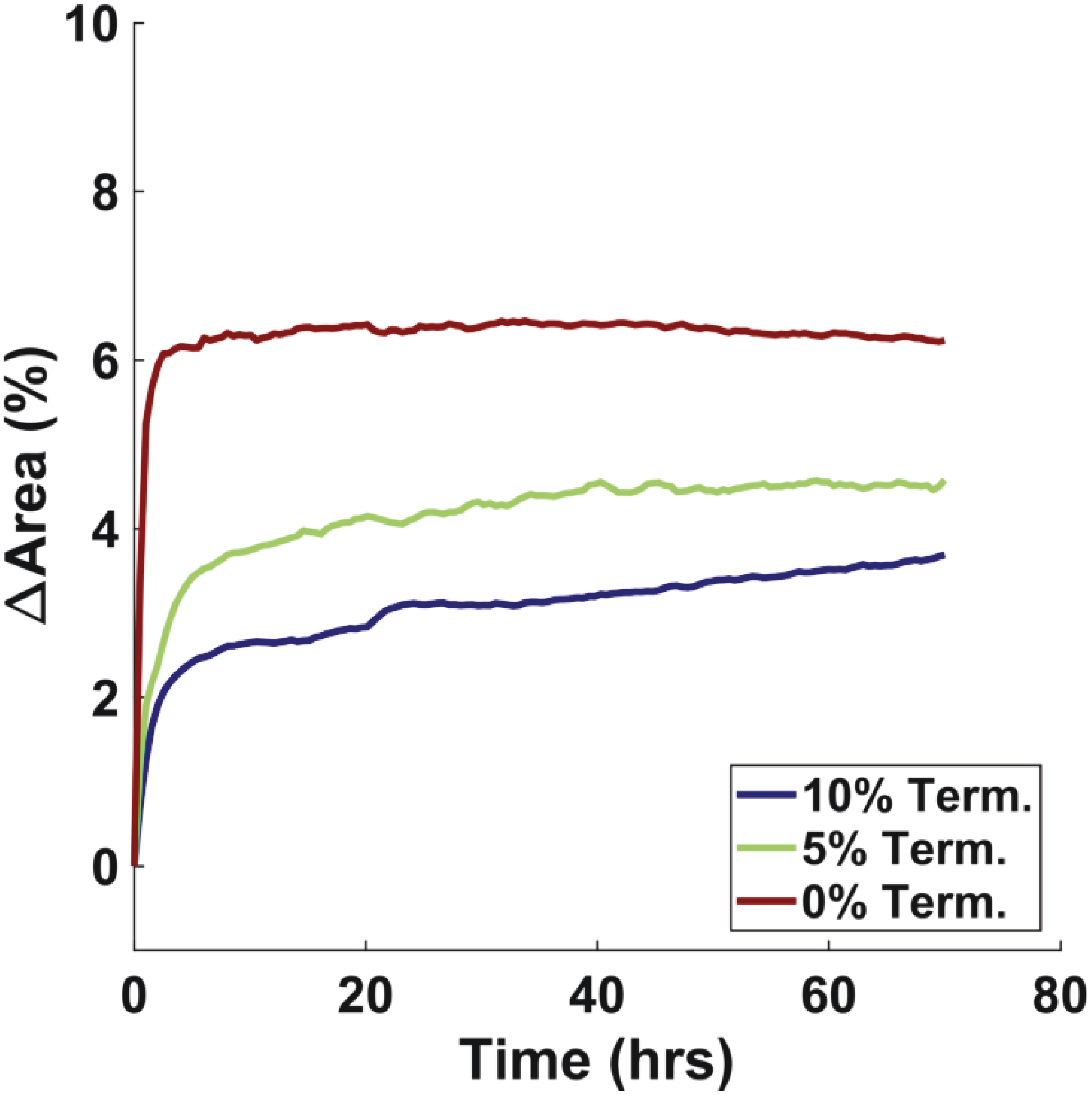


Figure S5.6: Locked crosslink hydrogel particles show minimal swelling with different percentages of hairpin terminator over 70 hours and 20 μM per hairpin type. Each curve is the average of two particles.

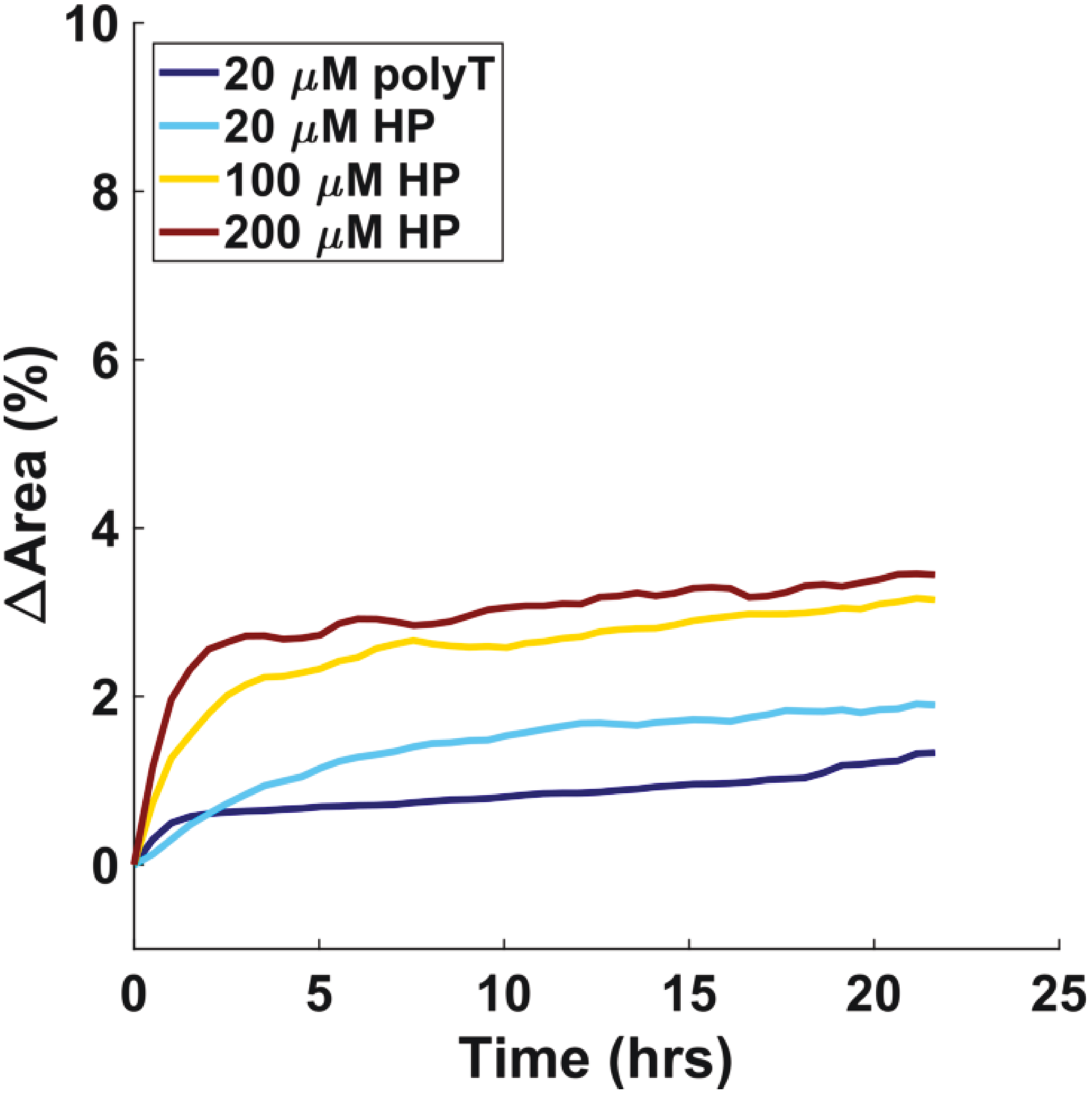


Figure S5.7: Swelling of locked particles incubated with different concentrations of hairpins. Particles showed only minimal swelling across all concentrations tested, indicating the robustness of the crosslink locking mechanism. In all cases, the fraction of hairpin terminator is 10%. When 20 μM of 20-mer polyT was added to particles, < 1% change in area was observed. Curves are the average of 2 – 4 particles.

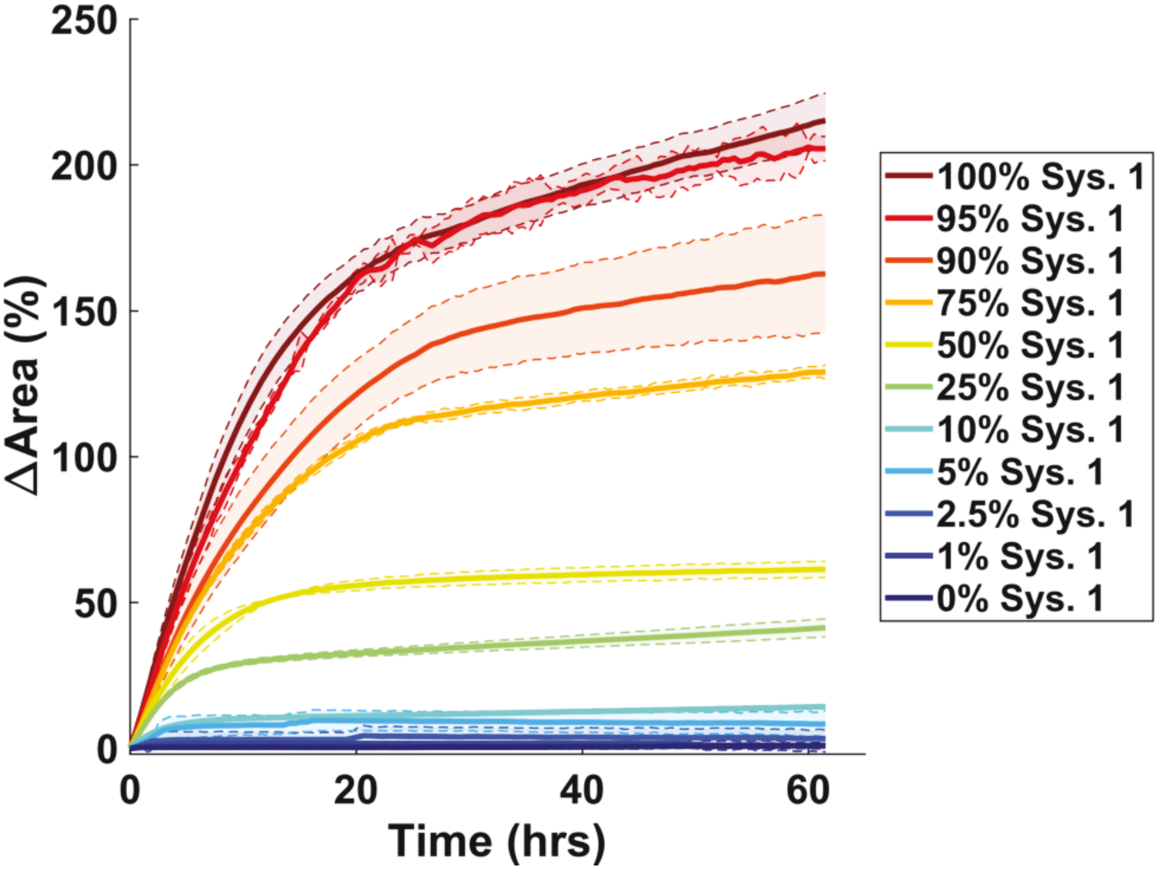


Figure S5.8: Kinetics of swelling when particles prepared with different fractions of expandable crosslinks *vs*. crosslinks with non-interacting sequences were incubated with hairpins. The crosslinks with alternative sequences were used to help understand how swelling would be affected if the concentration of key strands limited the number of locked crosslinks becoming unlocked. The sequences of the non-interacting crosslinks were designed to not interact with the expandable crosslinks or hairpins (Table S5.1). Hairpins with 10% terminator (see Figure S5.3) contained sequences to direct incorporation into the expandable crosslinks. The legend gives the percentage of crosslinks that are expandable *via* hairpins. Crosslinks that cannot expand limit the expansion capabilities of the hydrogel matrix, reducing overall swelling. Each curve is the average of 8 particles for the 100% Sys. 1 case and 3 particles for all other cases.

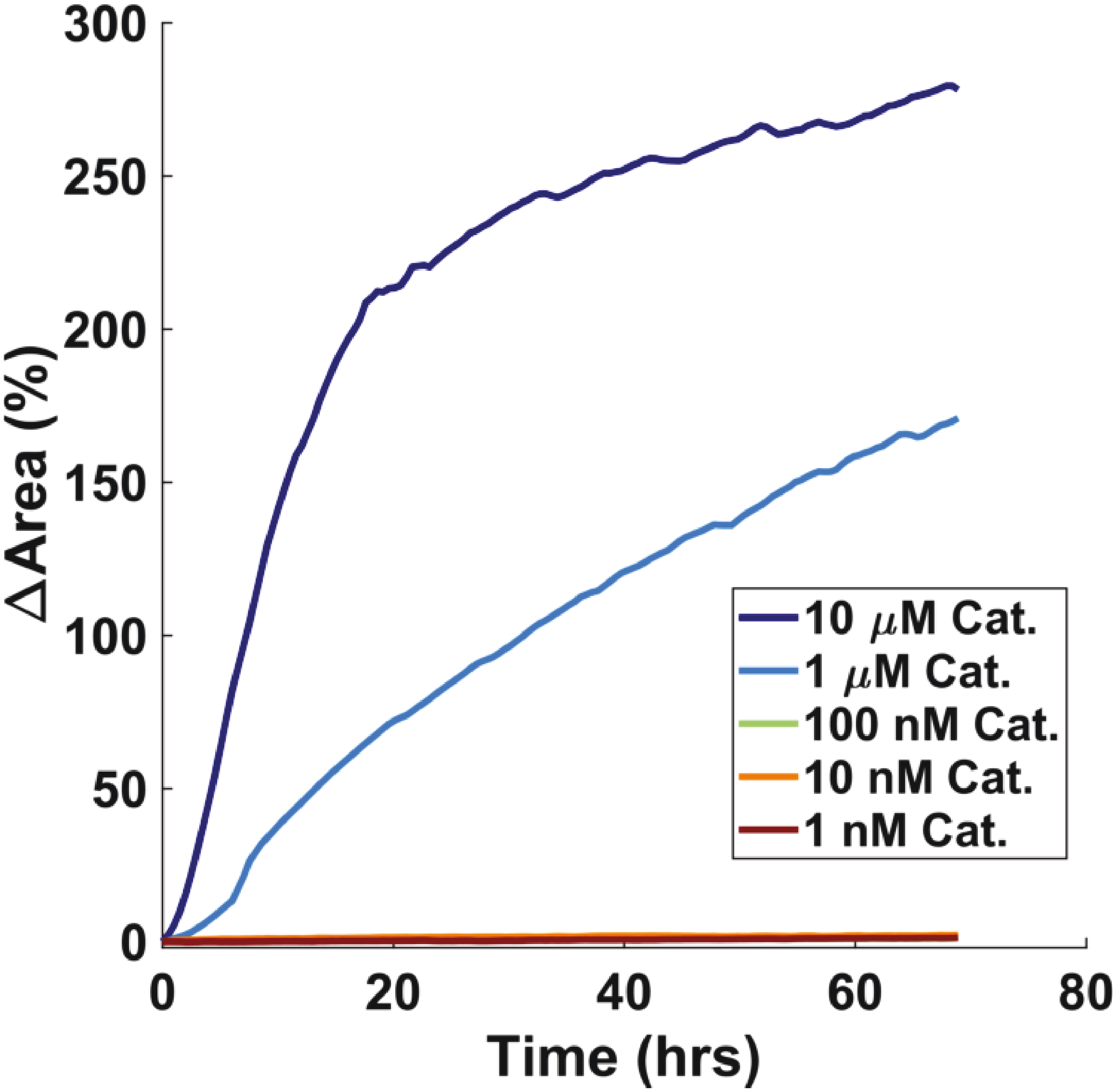


Figure S5.9: Swelling of locked particles with hairpin fuel and different concentrations of Catalyst strand. The swelling kinetics increases with increasing Catalyst concentration. Hairpins are 20 μM per type with 10% terminator. Curves for 1, 10, and 100 nM Catalyst are indistinguishable from each other. Curves are the average of two particles.

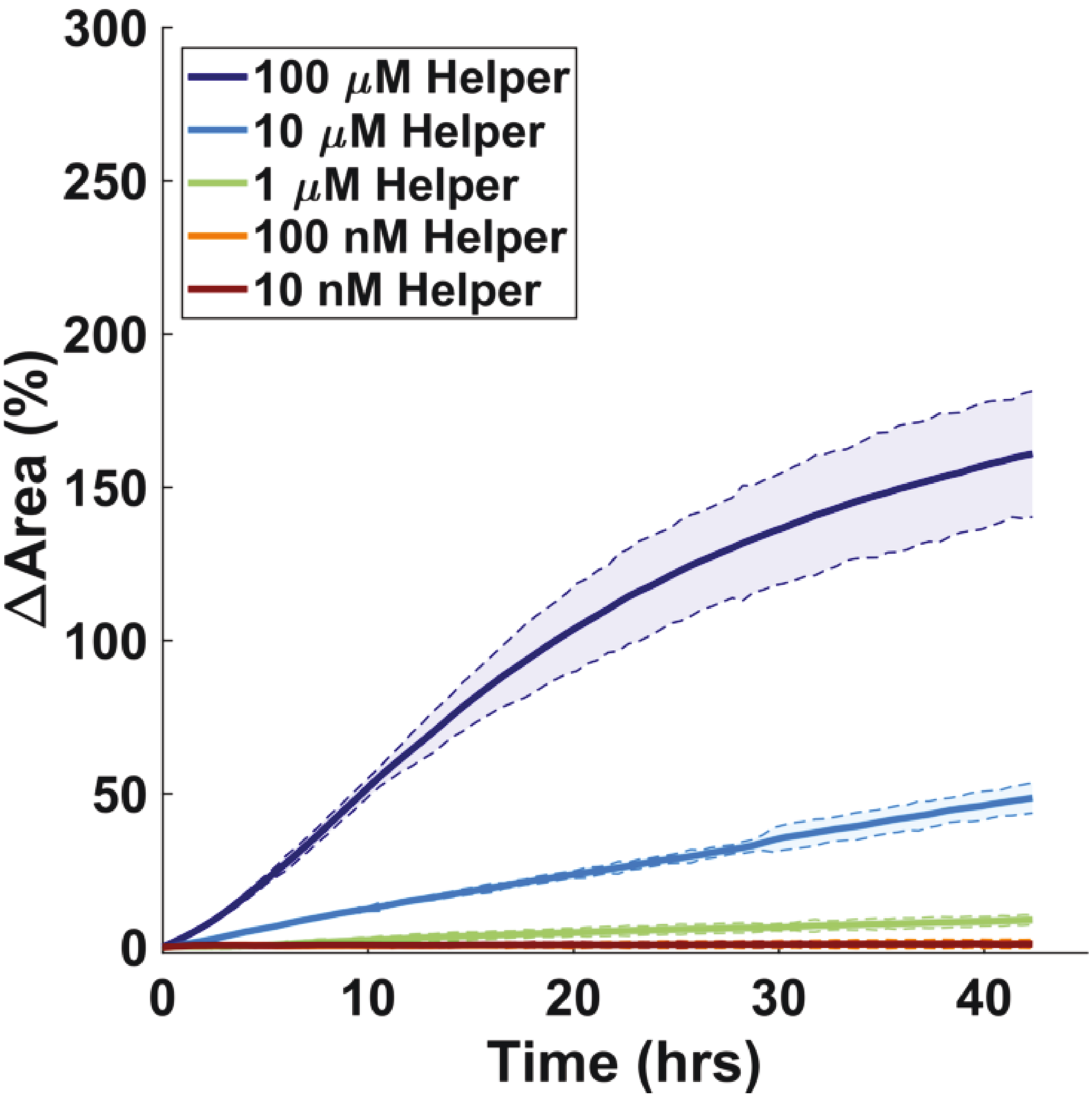


Figure S5.10: Swelling of locked particles in the presence of Helper strand and hairpin fuel. Particles were incubated with different concentrations of Helper strand and 20 μM hairpins with 10% terminator. In all cases, swelling was lower and slower than that of Catalyst with hairpins at the same concentrations. For all controller experiments, 10 μM Helper strand was chosen due to the balance of undesired swelling shown here, and the high degree of catalytic turnover when Helper and Catalyst were added together (Figure S5.11). Curves are the average of two particles.

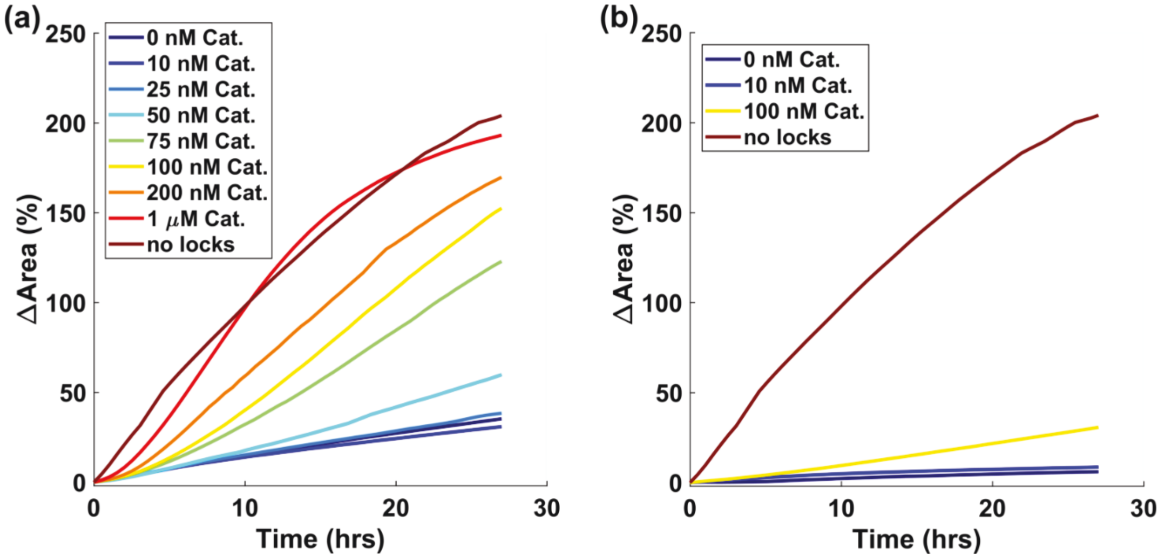


Figure S5.11: Comparison of swelling using both Catalyst and Helper strands at either 10 μM Helper (a) or 1 μM Helper (b) strands. When the concentration of the Helper strand is set to 1 μM, the degree of swelling is significantly limited over 25 hours compared to swelling with 10 μM Helper, indicating that a higher [Helper] is needed (*e.g.*, 10 μM) for high catalytic turnover and high-degrees of swelling. The swelling of particles prepared without locks is shown for comparison. The hairpin concentration for all curves is 20 μM per type with 10% terminator. Curves are the average of 2 – 6 particles.

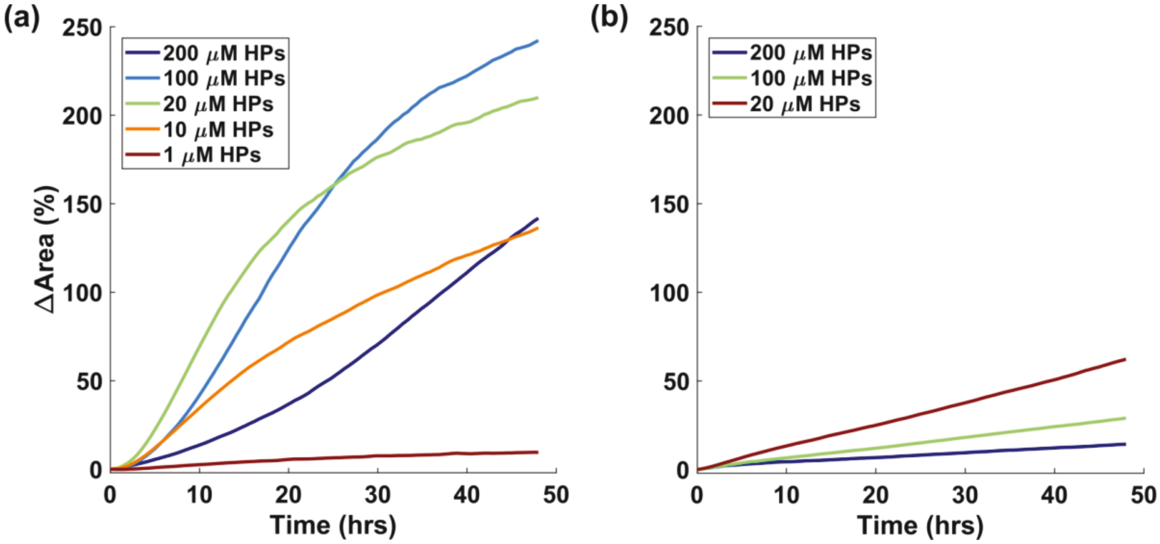


Figure S5.12: Swelling of particles with 100 nM (a) or 10 nM (b) Catalyst and 10 μM Helper strands incubated with varying concentrations of hairpin. Concentrations in the legends are "per hairpin type" and the percentage of hairpin terminator in all cases was fixed at 10%. The speed of swelling decreased at higher concentrations of hairpin, likely due to sequestration of Catalyst and/or Helper strands by the hairpins due to shared complimentary active site domains (Figures S5.3 and S5.19). At hairpin concentrations below 20 μM, the total number of hairpins available for incorporation and the rate of incorporation is decreased, leading to slower or very little swelling. Curves are single particles (200 μM) or an average of 2 – 4 particles.

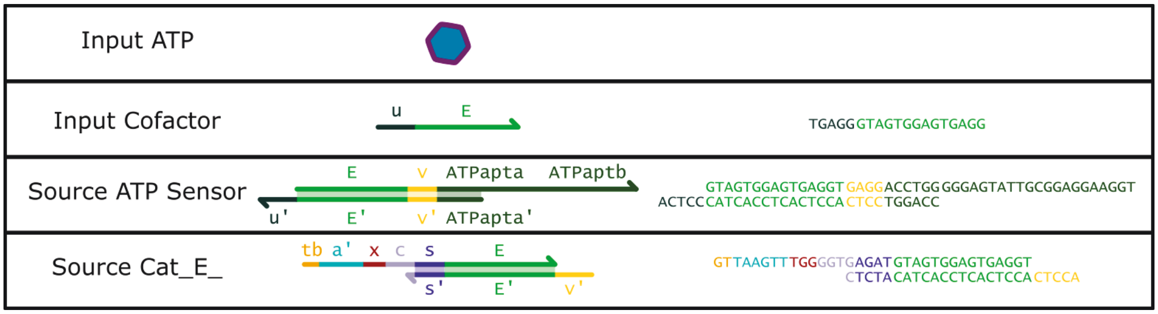


Figure S5.13: Schematic of the components of the ATP aptasensor circuit. *ATPapta* and *ATPaptb* (dark green domain) are the domains containing the ATP aptamer sequence. The toeholds *u* and *v* and domain *E* were designed using NUPACK.[116] Both ATP and the Cofactor strand are needed for ATP detection.

Figure S5.14: Reporter reaction to detect the presence of Catalyst strand or catalyst-sequence-containing strands using the fluorophore-quencher assay on a qPCR. The Catalyst strand binds *via* the toehold s plus one extra base (5 bases total) and branch migrates off the quencher strand, allowing for the FAM fluorophore to be detected.

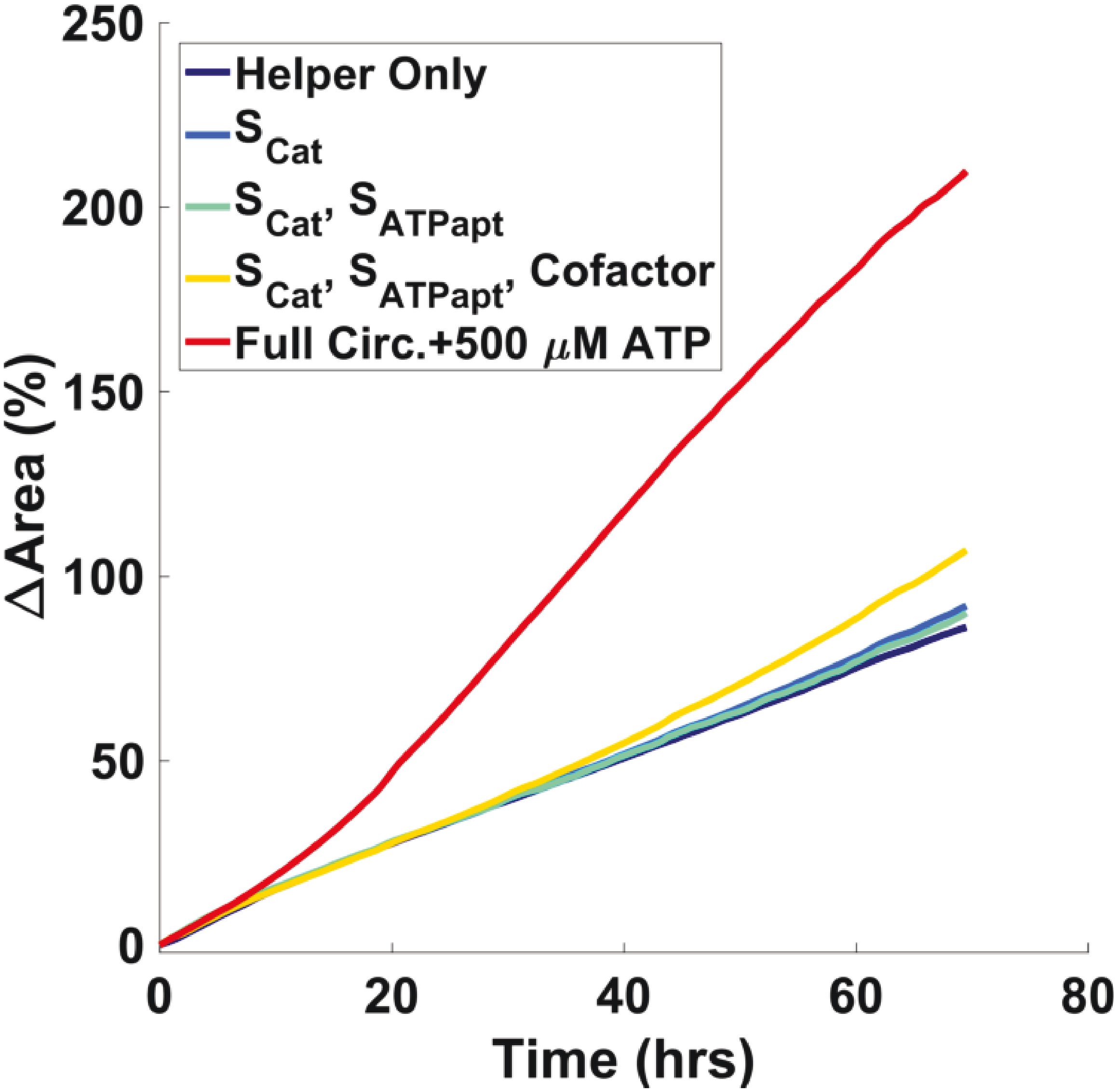


Figure S5.15: Effect of individual aptasensor circuit components (Figure S5.13) on particle swelling. Particles incubated with Source complexes (S_{cat} and S_{ATPapt}) did not show increased swelling over that of 10 μM Helper strand alone, indicating minimal to no undesired release of Catalyst in the absence of inputs. Over 70 hours, a small amount of undesired Catalyst strand is released when the Cofactor strand is added to the circuit in addition to the circuit complexes. The full circuit reacted with 500 μM ATP is shown for comparison. Curves are the average of 3 – 6 particles.

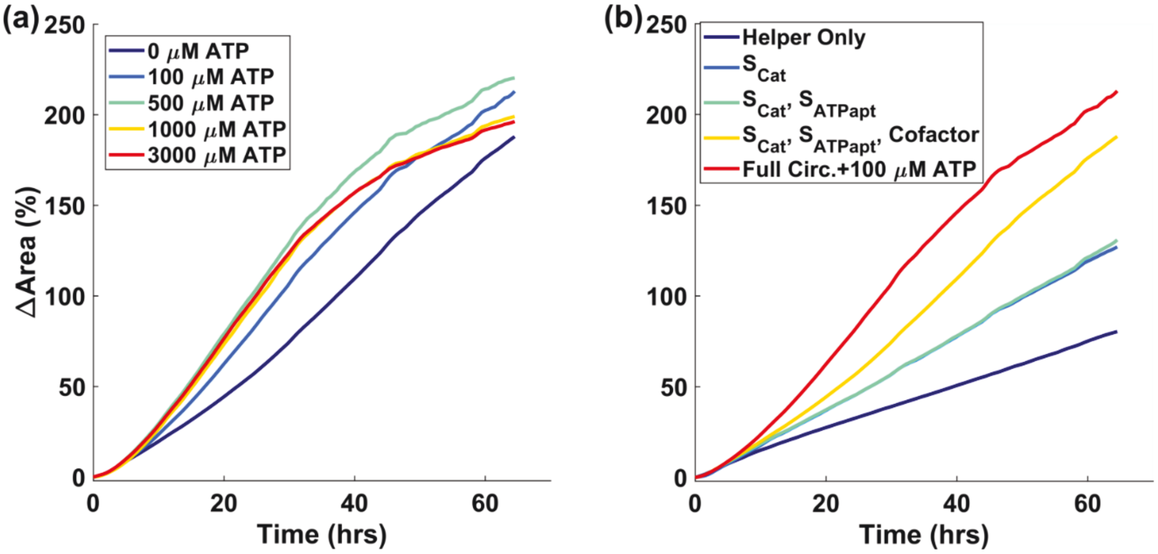


Figure S5.16: Increasing the baseline concentration of the aptasensor circuit components (Source complexes and Cofactor strand) to 200 nM decreases sensitivity to ATP concentration. Undesired reactions between circuit components and the hydrogel crosslinks in the absence of ATP releases enough Catalyst strand to initiate a substantial degree of swelling. This decreases sensitivity to ATP concentration and reduces the difference in swelling when ATP is present *vs*. absent. These undesired reactions were not found to be present in a significant fashion at 100 nM circuit components (Figure S5.15).

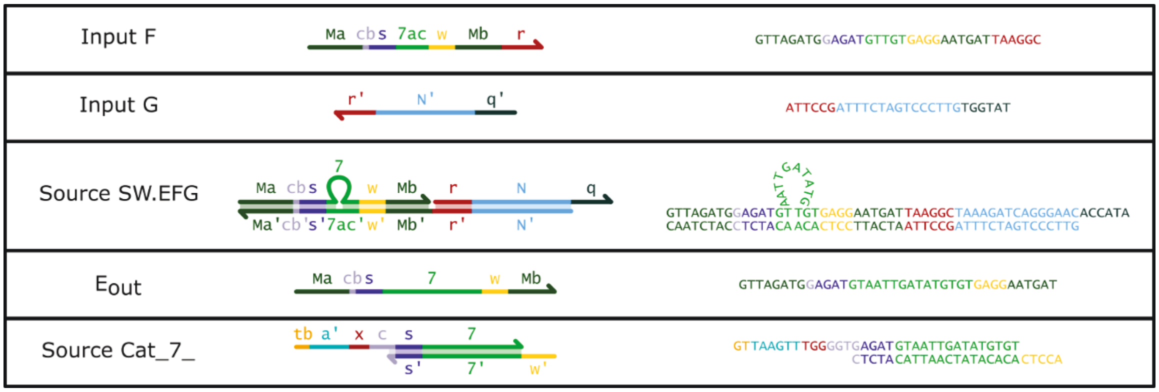


Figure S5.17: Schematic of the components of the Logic Circuit based on a design by Seelig *et al.*[107] The toehold, *w*, was designed using NUPACK[116] and the *7* domain sequence was taken from Qian and Winfree.[87]

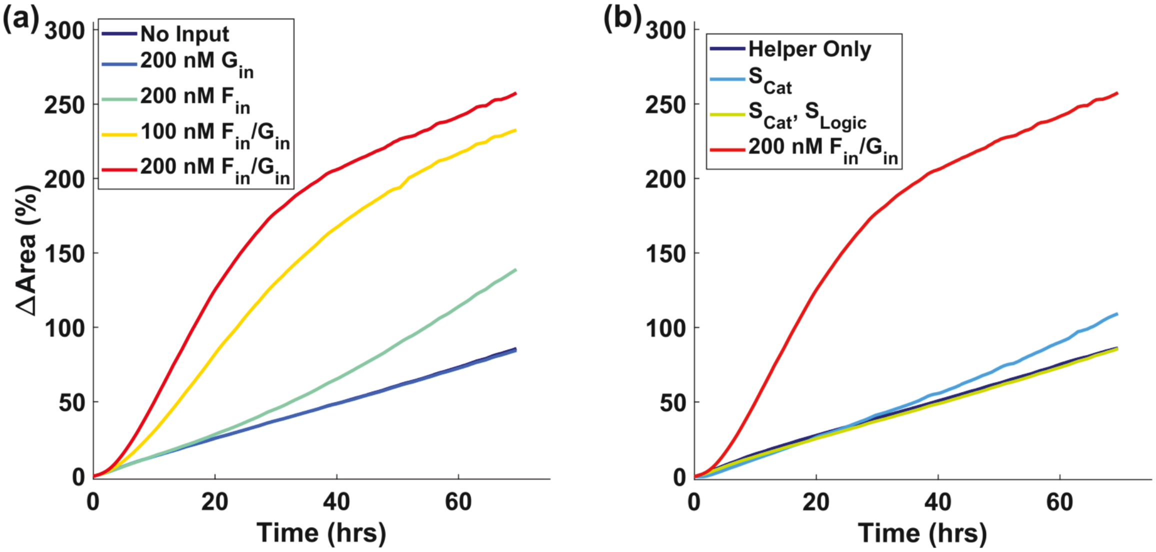


Figure S5.18: Swelling kinetics of the particles when incubated with the Logic Circuit over 70 hours. (a) The same experiment as in Figure 5.4 for 70 hours. The swelling induced from adding F_{in} could result from some undesired reactions that slowly release Catalyst over time as F_{in} contains the *s* and *cb* domains of the Catalyst strand (Figure S5.17). (b) Inclusion of the controller in the absence of inputs with locked particles and the amplification system did not significantly increase baseline swelling.

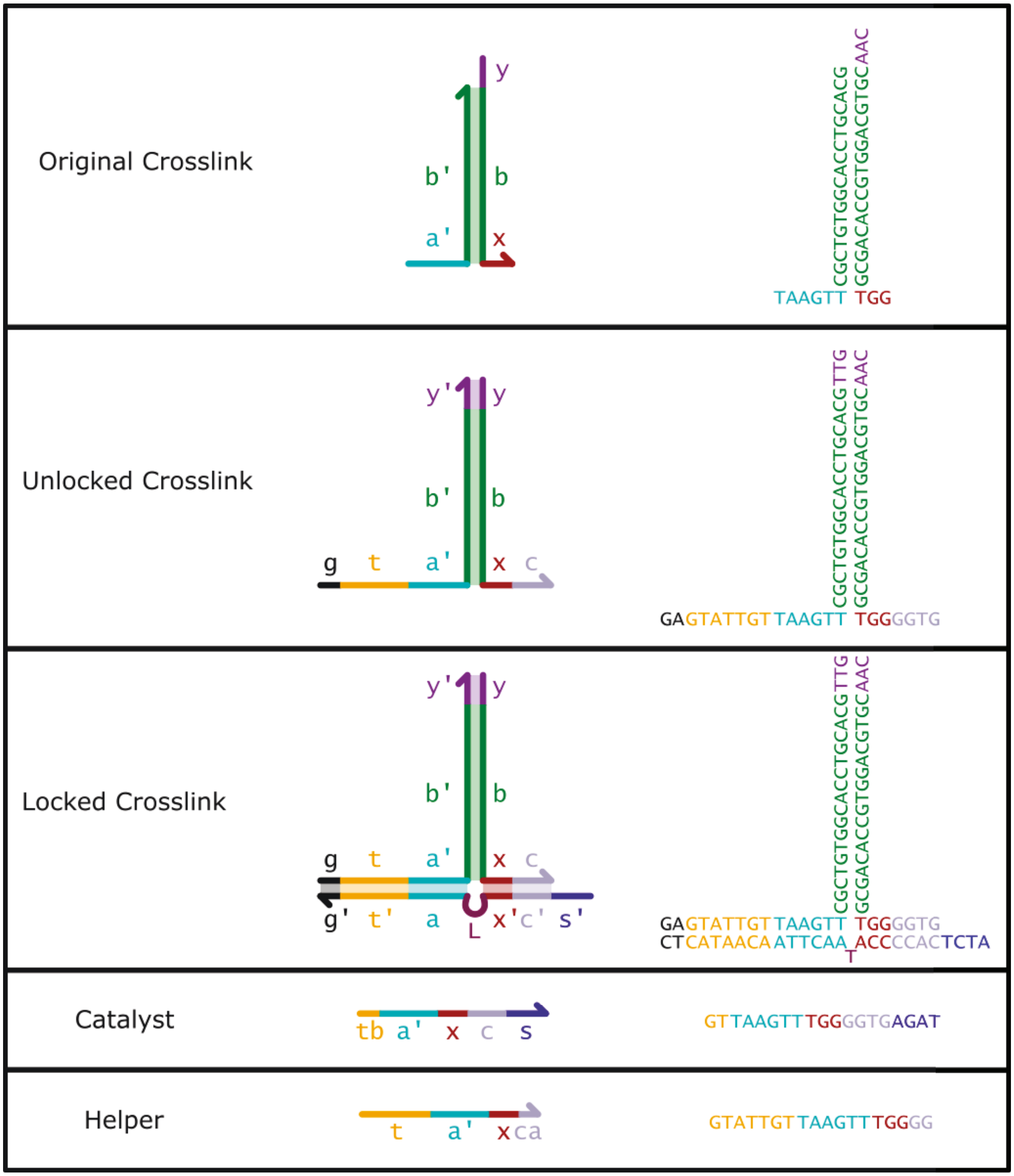


Figure S5.19: DNA crosslink sequences. The original crosslink design is also used in Cangialosi *et al.*[239] The additional sequences in the locked/unlocked crosslinks, Catalyst, and Helper strands were designed using NUPACK.[116] The domains *a'* and *x* (teal and red domains) represent the initial active site for hairpin fuel incorporation.

Table S5.1: List of sequences used in this study. Sequences were either taken from previous literature[87,107,239] or designed using NUPACK[116] as noted in Figures S5.3, S5.13, S5.17, and S5.19. The crosslinks *AlHPCC(v1)* and *RlHPCC(v1)* were used in Figure S5.4 and for measuring the swelling of the hydrogels prepared without locks (data labeled "no locks"). The experiment for Figure S5.5 was conducted using *A1* and *R1* crosslinks. The crosslinks *AlHPCC*, *RlHPCC*, *ASys2*, and *RSys2* were used for the particle swelling measurements presented in Figure S5.8.

Strand Name	**Role**	**Sequence**
		Crosslinks
A1	Original Crosslink	/5Acryd/TAAGTT CGCTGTGGCACCTGCACG
R1	Original Crosslink	/5Acryd/CAA CGTGCAGGTGCCACAGCG TGG
AlHPCC	Lockable Crosslink	/5Acryd/TT GA GTATTGT TAAGTT CGCTGTGGCACCTGCACG TTG
RlHPCC	Lockable Crosslink	/5Acryd/CAA CGTGCAGGTGCCACAGCG TGG GGTG TTT
AlHPCC (v1)	Crosslink (Fig. S5.4, "no locks" data)	/5Acryd/TGGT TAAGTT CGCTGTGGCACCTGCACG TTG
RlHPCC (v1)	Crosslink (Fig. S5.4, "no locks" data)	/5Acryd/CAA CGTGCAGGTGCCACAGCG TGG GG
ASys2	Sys. 2 Crosslink (Fig. S5.8)	/5Acryd/TT GT TATGTAT CTGTCT GCCTACCACTCCGTTGCG AAT
RSys2	Sys. 2 Crosslink (Fig. S5.8)	/5Acryd/ATT CGCAACGGAGTGGTAGGC TTT GA AT TTT
		Locking/Unlocking Strands
GblHPCC	Locking Strand	ATCT CACC CCA T AACTTA ACAATAC TC
FClHPCC	Key Strand	GA GTATTGT TAAGTT A TGG GGTG AGAT
HelperlHPCC	Helper Strand	GTATTGT TAAGTT TGG GG
CataylstlHPCC	Catalyst Strand	GT TAAGTT TGG GGTG AGAT
		Hairpin Strands
H1	Hairpin Monomer	CCA CGCTGTGGCACCTGCACG CACCCA CGTGCAGGTGCCACAGCG AACTTA
H2	Hairpin Monomer	TGGGTG CGTGCAGGTGCCACAGCG TAAGTT CGCTGTGGCACCTGCACG TTG
H1terminator	Hairpin Monomer	CCA CGCTGTGGCACCTGCACG TAGACT CGTGCAGGTGCCACAGCG AACTTA
H2terminator	Hairpin Monomer	TGGGTG CGTGCAGGTGCCACAGCG GCCTAG CGCTGTGGCACCTGCACG TTG

Logic Converter		
Eo7Cat	Logic Gate	GTTAGATG G AGAT GT AATTGATATG T GT GAG G AATGAT
GbEFG	Logic Gate	GTTCCCTGATCTTTA GCCTTA ATCATT C CTC AC A AC ATCT C CATCTAAC
G	Logic Gate	TAAGGC TAAAGATCAGGGAAC ACCATA
G.in	Logic Input	TATGGT GTTCCCTGATCTTTA GCCTTA
Fcatst.in	Logic Input/Purif.	GTTAGATG G AGAT GT T GT GAG G AATGAT TAAGGC
G.in.NoToe	Logic Gate Purification	GTTCCCTGATCTTTAGCCTTA
Wcatalyst_7	Logic-Cat. Converter	GT TAAGTT TGG GGTG AGAT GT AATTGATATGT GT
GbC7	Logic-Cat. Converter	AC CTC AC ACATATCAATT AC ATCT C
ATP Sensor/Converter		
Cof.tapt_eta	Cofactor	TGAGG GT AGTGGAGTGAG G
Weta_ATPapt	ATPSensor Gate	GT AGTGGAGTGAG GT GAG G ACCTGGGGGAGTATTGCGGAGGAAGGT
Gbeta_ATPapt	ATPSensor Gate	CCAGGT C CTC AC CTCACTCCACT AC CCTCA
Wcatalyst_eta	Aptamer-Cat. Converter	GT TAAGTT TGG GGTG AGAT GT AGTGGAGTGAG GT
GbCeta	Aptamer-Cat. Converter	AC CTC AC CTCACTCCACT AC ATCT C
Controller Reporting Assay Strands		
Rv(Wcat)q	Reporter	/5IABkFQ/GT TAAGTT TGG GG TG
Rb(Wcat)f	Reporter	C ATCT CA CC CCA AACTTA AC/36-FAM/
PolyT20	polyT	TTTTTTTTTTTTTTTTTTTT

5.6 Acknowledgements

The authors thank D. Scalise, A. Cangialosi, and S. Schaffter for thoughtful discussions. This research is supported by the Department of Energy Early Career Award.

CHAPTER 6: CONCLUSION

One could envision using the concepts and technologies demonstrated in *Chapters 3 – 5* to build more complex soft-material devices or self-assembled materials. These devices could be engineered with embedded aptamers or other sensing modules such that the signal activation is localized to a specific hydrogel region compared to a well-mixed solution.[112,114] Indeed, soft robots or smart materials will necessarily require this level of integration in order to function in physiological or other environments where the efficiency of the non-integrated material is impaired by the diffusion of controller circuit components away from the material or cleared from the local system.

Additionally, the hydrogel crosslink design used in *Chapter 5* could be modified to incorporate force-sensing capabilities using DNA hairpins with force-sensitive zipper or shear-mode motifs.[235] The force-sensing modules would enable the engineering of feedback circuits where forces enacted on the material are translated into expansion of the crosslinks local to the source of the force, and would direct forces back onto the polymer network and translocation of mechanical signals across the network. Force-translational circuits could also be designed where strain applied to the DNA-crosslinked material directs the release of a biomolecule or directs actuation such as bilayer curling. Alongside these experiments, quantitative models will need to be developed to capture the complex chemo-mechanical mechanisms occurring within the DNA-crosslinked hydrogels to more efficiently engineer and design such systems. The models will not only require equations to predict the effect of strain/stress on the polymer network,[253,254] but also how reaction-diffusion strand-displacement reactions[96,115] can control material properties on the crosslinker level.[156]

As the price of DNA synthesis with or without modifications decreases, the breadth and depth of the technologies centered around DNA controllers, circuits, sensors, and nanostructures will continue to grow. In a relatively short time, the material-based potential of

DNA has expanded from small 4 strand structures in the 1980's,[3] to self-contained structures with 100's of strands[14,45] and even larger DNA-integrated materials over a millimeter in size. I expect with the demonstrated experimental, theoretical, and computational foundation that has been established thus far, the field of DNA nanotechnology will continue to investigate applicable technologies in such directions as microscopy, biology, material mechanics, and in computation where the self-assemblies and DNA strand-displacement systems currently developed will just be small components of multi-functional, complex systems.

BIBLIOGRAPHY

1. Seeman, N. C. Nucleic acid junctions and lattices. *J. Theor. Biol.* **99,** 237–247 (1982).
2. Seeman, N. C. & Kallenbach, N. R. Design of immobile nucleic acid junctions. *Biophys. J.* **44,** 201–209 (1983).
3. Kallenbach, N. R., Ma, R.-I. & Seeman, N. C. An immobile nucleic acid junction constructed from oligonucleotides. *Nature* **305,** 829–831 (1983).
4. Chen, J. H. & Seeman, N. C. Synthesis from DNA of a molecule with the connectivity of a cube. *Nature* **350,** 631–633 (1991).
5. Seeman, N. C. DNA engineering and its application to nanotechnology. *Trends Biotechnol.* **17,** 437–443 (1999).
6. Carlson, R. The changing economics of DNA synthesis. *Nat. Biotechnol.* **27,** 1091–1094 (2009).
7. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).
8. Ke, Y., Ong, L. L., Shih, W. M. & Yin, P. Three-dimensional structures self-assembled from DNA bricks. *Science* **338,** 1177–1183 (2012).
9. Yang, Y., Han, D., Nangreave, J., Liu, Y. & Yan, H. DNA origami with double-stranded DNA as a unified scaffold. *ACS Nano* **6,** 8209–8215 (2012).
10. Douglas, S. M. *et al.* Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res.* **37,** 5001–5006 (2009).
11. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325,** 725–730 (2009).
12. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459,** 414–418 (2009).

13. Ke, Y. *et al.* Multilayer DNA origami packed on a square lattice. *J. Am. Chem. Soc.* **131,** 15903–15908 (2009).

14. Ke, Y. *et al.* DNA brick crystals with prescribed depths. *Nat. Chem.* **6,** 994–1002 (2014).

15. Zhao, Z., Yan, H. & Liu, Y. A route to scale up DNA origami using DNA tiles as folding staples. *Angew. Chem. Int. Ed Engl.* **49,** 1414–1417 (2010).

16. Wei, B., Dai, M. & Yin, P. Complex shapes self-assembled from single-stranded DNA tiles. *Nature* **485,** 623–626 (2012).

17. Knudsen, J. B. *et al.* Routing of individual polymers in designed patterns. *Nat. Nanotechnol.* **10,** 892–898 (2015).

18. Krissanaprasit, A. *et al.* Programmed Switching of Single Polymer Conformation on DNA Origami. *ACS Nano* **10,** 2243–2250 (2016).

19. Tokura, Y. *et al.* Bottom-Up Fabrication of Nanopatterned Polymers on DNA Origami by In Situ Atom-Transfer Radical Polymerization. *Angew. Chem. Int. Ed Engl.* **55,** 5692–5697 (2016).

20. Schmied, J. J. *et al.* DNA origami-based standards for quantitative fluorescence microscopy. *Nat. Protoc.* **9,** 1367–1391 (2014).

21. Steinhauer, C., Jungmann, R., Sobey, T. L., Simmel, F. C. & Tinnefeld, P. DNA origami as a nanoscopic ruler for super-resolution microscopy. *Angew. Chem. Int. Ed Engl.* **48,** 8870–8873 (2009).

22. Liu, Q., Song, C., Wang, Z.-G., Li, N. & Ding, B. Precise organization of metal nanoparticles on DNA origami template. *Methods San Diego Calif* **67,** 205–214 (2014).

23. Kuzyk, A., Laitinen, K. T. & Törmä, P. DNA origami as a nanoscale template for protein assembly. *Nanotechnology* **20,** 235305 (2009).

24. Wilner, O. I. *et al.* Enzyme cascades activated on topologically programmed DNA scaffolds. *Nat. Nanotechnol.* **4,** 249–254 (2009).

25. Fu, J. *et al.* Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* **9,** 531–536 (2014).

26. Linko, V., Eerikäinen, M. & Kostiainen, M. A. A modular DNA origami-based enzyme cascade nanoreactor. *Chem. Commun. Camb. Engl.* **51,** 5351–5354 (2015).

27. Schüller, V. J. *et al.* Cellular Immunostimulation by CpG-Sequence-Coated DNA Origami Structures. *ACS Nano* **5,** 9696–9702 (2011).

28. Shaw, A. *et al.* Spatial control of membrane receptor function using ligand nanocalipers. *Nat. Methods* **11,** 841–846 (2014).

29. Sellner, S. *et al.* DNA nanotubes as intracellular delivery vehicles in vivo. *Biomaterials* **53,** 453–463 (2015).

30. Angelin, A. *et al.* Multiscale Origami Structures as Interface for Cells. *Angew. Chem. Int. Ed Engl.* **54,** 15813–15817 (2015).

31. Mikkilä, J. *et al.* Virus-encapsulated DNA origami nanostructures for cellular delivery. *Nano Lett.* **14,** 2196–2200 (2014).

32. Ora, A. *et al.* Cellular delivery of enzyme-loaded DNA origami. *Chem. Commun. Camb. Engl.* **52,** 14161–14164 (2016).

33. Kuzyk, A. *et al.* DNA-based self-assembly of chiral plasmonic nanostructures with tailored optical response. *Nature* **483,** 311–314 (2012).

34. Ding, B. *et al.* Gold Nanoparticle Self-Similar Chain Structure Organized by DNA Origami. *J. Am. Chem. Soc.* **132,** 3248–3249 (2010).

35. Tan, S. J., Campolongo, M. J., Luo, D. & Cheng, W. Building plasmonic nanostructures with DNA. *Nat. Nanotechnol.* **6,** 268–276 (2011).

36. Chao, J., Lin, Y., Liu, H., Wang, L. & Fan, C. DNA-based plasmonic nanostructures. *Mater. Today* **18,** 326–335 (2015).

37. Hung, A. M. *et al.* Large-area spatially ordered arrays of gold nanoparticles directed by lithographically confined DNA origami. *Nat. Nanotechnol.* **5,** 121–126 (2010).

38. Gállego, I. *et al.* DNA-Origami-Driven Lithography for Patterning on Gold Surfaces with Sub-10 nm Resolution. *Adv. Mater. Deerfield Beach Fla* **29,** (2017).

39. Busuttil, K., Rotaru, A., Dong, M., Besenbacher, F. & Gothelf, K. V. Transfer of a protein pattern from self-assembled DNA origami to a functionalized substrate. *Chem. Commun. Camb. Engl.* **49,** 1927–1929 (2013).

40. Schreiber, R. *et al.* DNA origami-templated growth of arbitrarily shaped metal nanoparticles. *Small Weinh. Bergstr. Ger.* **7,** 1795–1799 (2011).

41. Helmi, S., Ziegler, C., Kauert, D. J. & Seidel, R. Shape-controlled synthesis of gold nanostructures using DNA origami molds. *Nano Lett.* **14,** 6693–6698 (2014).

42. Sun, W. *et al.* Casting inorganic structures with DNA molds. *Science* **346,** 1258361 (2014).

43. Fern, J., Lu, J. & Schulman, R. The Energy Landscape for the Self-Assembly of a Two-Dimensional DNA Origami Complex. *ACS Nano* **10,** 1836–1844 (2016).

44. Zenk, J., Tuntivate, C. & Schulman, R. Kinetics and Thermodynamics of Watson-Crick Base Pairing Driven DNA Origami Dimerization. *J. Am. Chem. Soc.* **138,** 3346–3354 (2016).

45. Marchi, A. N., Saaem, I., Vogen, B. N., Brown, S. & LaBean, T. H. Toward Larger DNA Origami. *Nano Lett.* **14,** 5740–5747 (2014).

46. Newcomb, W. W. *et al.* Assembly of the Herpes Simplex Virus Capsid: Characterization of Intermediates Observed During Cell-free Capsid Formation. *J. Mol. Biol.* **263,** 432–446 (1996).

47. Katen, S. & Zlotnick, A. Chapter 14 The Thermodynamics of Virus Capsid Assembly. in *Methods in Enzymology* **455,** 395–417 (Elsevier, 2009).

48. Hagan, M. F. & Chandler, D. Dynamic pathways for viral capsid assembly. *Biophys. J.* **91,** 42–54 (2006).

49. Woodson, S. A. RNA folding pathways and the self-assembly of ribosomes. *Acc. Chem. Res.* **44,** 1312–1319 (2011).

50. Alberts, B. *et al.* The Self-Assembly and Dynamic Structure of Cytoskeletal Filaments. in *Molecular Biology of the Cell* (Garland Science, 2002).

51. Marsh, J. A. & Teichmann, S. A. Structure, dynamics, assembly, and evolution of protein complexes. *Annu. Rev. Biochem.* **84,** 551–575 (2015).

52. Halley, J. D. & Winkler, D. A. Consistent concepts of self-organization and self-assembly. *Complexity* **14,** 10–17 (2008).

53. Turner, D. A., Baillie-Johnson, P. & Martinez Arias, A. Organoids and the genetically encoded self-assembly of embryonic stem cells. *BioEssays News Rev. Mol. Cell. Dev. Biol.* **38,** 181–191 (2016).

54. Couzin, I. D. & Krause, J. Self-Organization and Collective Behavior in Vertebrates. in *Advances in the Study of Behavior* **32,** 1–75 (Elsevier, 2003).

55. Mandal, D., Nasrolahi Shirazi, A. & Parang, K. Self-assembly of peptides to nanostructures. *Org. Biomol. Chem.* **12,** 3544–3561 (2014).

56. Hu, Y. *et al.* Spatiotemporal control of the creation and immolation of peptide assemblies. *Coord. Chem. Rev.* **320–321,** 2–17 (2016).

57. Ferrer-Miralles, N. *et al.* Engineering protein self-assembling in protein-based nanomedicines for drug delivery and gene therapy. *Crit. Rev. Biotechnol.* **35,** 209–221 (2015).

58. Mout, R. *et al.* Programmed Self-Assembly of Hierarchical Nanostructures through Protein-Nanoparticle Coengineering. *ACS Nano* (2017). doi:10.1021/acsnano.6b07258

59. Mao, C., Sun, W. & Seeman, N. C. Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy. *J. Am. Chem. Soc.* **121,** 5437–5443 (1999).

60. Li, Z., Wang, L., Yan, H. & Liu, Y. Effect of DNA Hairpin Loops on the Twist of Planar DNA Origami Tiles. *Langmuir* **28,** 1959–1965 (2012).

61. He, Y., Tian, Y., Ribbe, A. E. & Mao, C. Highly connected two-dimensional crystals of DNA six-point-stars. *J. Am. Chem. Soc.* **128,** 15978–15979 (2006).

62. He, Y. *et al.* Sequence symmetry as a tool for designing DNA nanostructures. *Angew. Chem. Int. Ed Engl.* **44,** 6694–6696 (2005).

63. Wang, P. *et al.* Programming Self-Assembly of DNA Origami Honeycomb Two-Dimensional Lattices and Plasmonic Metamaterials. *J. Am. Chem. Soc.* **138,** 7733–7740 (2016).

64. Park, S. H. *et al.* Three-helix bundle DNA tiles self-assemble into 2D lattice or 1D templates for silver nanowires. *Nano Lett.* **5,** 693–696 (2005).

65. Zenk, J. & Schulman, R. An Assembly Funnel Makes Biomolecular Complex Assembly Efficient. *PLoS ONE* **9,** e111233 (2014).

66. Marsh, J. A. *et al.* Protein Complexes Are under Evolutionary Selection to Assemble via Ordered Pathways. *Cell* **153,** 461–470 (2013).

67. Hedges, L. O., Mannige, R. V. & Whitelam, S. Growth of equilibrium structures built from a large number of distinct component types. *Soft Matter* **10,** 6404 (2014).

68. Murugan, A., Zou, J. & Brenner, M. P. Undesired usage and the robust self-assembly of heterogeneous structures. *Nat. Commun.* **6,** 6203 (2015).

69. Jacobs, W. M., Reinhardt, A. & Frenkel, D. Rational design of self-assembly pathways for complex multicomponent structures. *Proc. Natl. Acad. Sci.* **112,** 6313–6318 (2015).

70. Wang, W. *et al.* Self-assembly of fully addressable DNA nanostructures from double crossover tiles. *Nucleic Acids Res.* **44,** 7989–7996 (2016).

71. Sharma, J., Chhabra, R., Liu, Y., Ke, Y. & Yan, H. DNA-templated self-assembly of two-dimensional and periodical gold nanoparticle arrays. *Angew. Chem. Int. Ed Engl.* **45,** 730–735 (2006).

72. Liu, W., Zhong, H., Wang, R. & Seeman, N. C. Crystalline Two-Dimensional DNA-Origami Arrays. *Angew. Chem. Int. Ed.* **50,** 264–267 (2011).

73. Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. Design and self-assembly of two-dimensional DNA crystals. *Nature* **394,** 539–544 (1998).

74. Rothemund, P. W. K., Papadakis, N. & Winfree, E. Algorithmic Self-Assembly of DNA Sierpinski Triangles. *PLoS Biol.* **2,** e424 (2004).

75. Barish, R. D., Schulman, R., Rothemund, P. W. K. & Winfree, E. An information-bearing seed for nucleating algorithmic self-assembly. *Proc. Natl. Acad. Sci.* **106,** 6054–6059 (2009).

76. Fujibayashi, K., Hariadi, R., Park, S. H., Winfree, E. & Murata, S. Toward Reliable Algorithmic Self-Assembly of DNA Tiles: A Fixed-Width Cellular Automaton Pattern. *Nano Lett.* **8,** 1791–1797 (2008).

77. Rajendran, A., Endo, M., Katsuda, Y., Hidaka, K. & Sugiyama, H. Programmed Two-Dimensional Self-Assembly of Multiple DNA Origami Jigsaw Pieces. *ACS Nano* **5,** 665–671 (2011).

78. Woo, S. & Rothemund, P. W. K. Programmable molecular recognition based on the geometry of DNA nanostructures. *Nat. Chem.* **3,** 620–627 (2011).

79. Kim, K. N., Sarveswaran, K., Mark, L. & Lieberman, M. Comparison of methods for orienting and aligning DNA origami. *Soft Matter* **7,** 4636 (2011).

80. Gerling, T., Wagenbauer, K. F., Neuner, A. M. & Dietz, H. Dynamic DNA devices and assemblies formed by shape-complementary, non-base pairing 3D components. *Science* **347,** 1446–1452 (2015).

81. Tikhomirov, G., Petersen, P. & Qian, L. Programmable disorder in random DNA tilings. *Nat. Nanotechnol.* **12,** 251–259 (2017).

82. Kim, D.-N., Kilchherr, F., Dietz, H. & Bathe, M. Quantitative prediction of 3D solution shape and flexibility of nucleic acid nanostructures. *Nucleic Acids Res.* **40,** 2862–2868 (2012).

83. Nangreave, J., Yan, H. & Liu, Y. Studies of Thermal Stability of Multivalent DNA Hybridization in a Nanostructured System. *Biophys. J.* **97,** 563–571 (2009).

84. Sedeh, R. S. *et al.* Computing Nonequilibrium Conformational Dynamics of Structured Nucleic Acid Assemblies. *J. Chem. Theory Comput.* **12,** 261–273 (2016).

85. Pan, K. *et al.* Lattice-free prediction of three-dimensional structure of programmed DNA assemblies. *Nat. Commun.* **5,** 5578 (2014).

86. Kholodenko, B. N. Cell-signalling dynamics in time and space. *Nat. Rev. Mol. Cell Biol.* **7,** 165–176 (2006).

87. Qian, L. & Winfree, E. Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **332,** 1196–1201 (2011).

88. Han, K., Liang, Z. & Zhou, N. Design strategies for aptamer-based biosensors. *Sensors* **10,** 4541–4557 (2010).

89. Meng, H.-M. *et al.* Aptamer-integrated DNA nanostructures for biosensing, bioimaging and cancer therapy. *Chem. Soc. Rev.* **45,** 2583–2602 (2016).

90. Yao, D. *et al.* Integrating DNA-Strand-Displacement Circuitry with Self-Assembly of Spherical Nucleic Acids. *J. Am. Chem. Soc.* **137,** 14107–14113 (2015).

91. Zhang, D. Y., Hariadi, R. F., Choi, H. M. T. & Winfree, E. Integrating DNA strand-displacement circuitry with DNA tile self-assembly. *Nat. Commun.* **4,** (2013).

92. Liang, H. *et al.* Functional DNA-containing nanomaterials: cellular applications in biosensing, imaging, and targeted therapy. *Acc. Chem. Res.* **47,** 1891–1901 (2014).

93. Jung, C. & Ellington, A. D. Diagnostic applications of nucleic acid circuits. *Acc. Chem. Res.* **47,** 1825–1835 (2014).

94. Khodakov, D. A., Khodakova, A. S., Huang, D. M., Linacre, A. & Ellis, A. V. Protected DNA strand displacement for enhanced single nucleotide discrimination in double-stranded DNA. *Sci. Rep.* **5,** 8721 (2015).

95. Zhang, D. Y., Chen, S. X. & Yin, P. Optimizing the specificity of nucleic acid hybridization. *Nat. Chem.* **4,** 208–214 (2012).

96. Zhang, D. Y. & Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J. Am. Chem. Soc.* **131,** 17303–17314 (2009).

97. Qian, L., Winfree, E. & Bruck, J. Neural network computation with DNA strand displacement cascades. *Nature* **475,** 368–372 (2011).

98. Srinivas, N. *et al.* On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Res.* **41,** 10641–10658 (2013).

99. Thachuk, C., Winfree, E. & Soloveichik, D. Leakless DNA Strand Displacement Systems. in *DNA Computing and Molecular Programming* (eds. Phillips, A. & Yin, P.) **9211,** 133–153 (Springer International Publishing, 2015).

100. Zhang, D. Y. & Seelig, G. Dynamic DNA nanotechnology using strand-displacement reactions. *Nat. Chem.* **3,** 103–113 (2011).

101. Fern, J. *et al.* DNA Strand-Displacement Timer Circuits. *ACS Synth. Biol.* **6,** 190–193 (2017).

102. Ouldridge, T. E., Sulc, P., Romano, F., Doye, J. P. K. & Louis, A. A. DNA hybridization kinetics: zippering, internal displacement and sequence dependence. *Nucleic Acids Res.* **41,** 8886–8895 (2013).

103. Šulc, P. *et al.* Sequence-dependent thermodynamics of a coarse-grained DNA model. *J. Chem. Phys.* **137,** 135101 (2012).

104. Machinek, R. R. F., Ouldridge, T. E., Haley, N. E. C., Bath, J. & Turberfield, A. J. Programmable energy landscapes for kinetic control of DNA strand displacement. *Nat. Commun.* **5,** 5324 (2014).

105. Broadwater, D. W. B. & Kim, H. D. The Effect of Basepair Mismatch on DNA Strand Displacement. *Biophys. J.* **110,** 1476–1484 (2016).

106. Olson, X., Kotani, S., Yurke, B., Graugnard, E. & Hughes, W. L. Kinetics of DNA Strand Displacement Systems with Locked Nucleic Acids. *J. Phys. Chem. B* **121,** 2594–2602 (2017).

107. Seelig, G., Soloveichik, D., Zhang, D. Y. & Winfree, E. Enzyme-free nucleic acid logic circuits. *Science* **314,** 1585–1588 (2006).

108. Genot, A. J., Bath, J. & Turberfield, A. J. Reversible Logic Circuits Made of DNA. *J. Am. Chem. Soc.* **133,** 20080–20083 (2011).

109. Zhang, D. Y., Turberfield, A. J., Yurke, B. & Winfree, E. Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **318,** 1121–1125 (2007).

110. Chen, X., Briggs, N., McLain, J. R. & Ellington, A. D. Stacking nonenzymatic circuits for high signal gain. *Proc. Natl. Acad. Sci.* **110,** 5386–5391 (2013).

111. Seelig, G., Yurke, B. & Winfree, E. Catalyzed Relaxation of a Metastable DNA Fuel. *J. Am. Chem. Soc.* **128,** 12211–12220 (2006).

112. Chirieleison, S. M., Allen, P. B., Simpson, Z. B., Ellington, A. D. & Chen, X. Pattern transformation with DNA circuits. *Nat. Chem.* **5,** 1000–1005 (2013).

113. Zadorin, A. S., Rondelez, Y., Galas, J.-C. & Estevez-Torres, A. Synthesis of programmable reaction-diffusion fronts using DNA catalyzers. *Phys. Rev. Lett.* **114,** 068301 (2015).

114. Scalise, D. & Schulman, R. Designing modular reaction-diffusion programs for complex pattern formation. *TECHNOLOGY* **02,** 55–66 (2014).

115. Zenk, J. *et al.* Stable DNA-based reaction–diffusion patterns. *RSC Adv.* **7,** 18032–18040 (2017).

116. Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* **32,** 170–173 (2011).

117. Grun, C., Werfel, J., Zhang, D. Y. & Yin, P. DyNAMiC Workbench: an integrated development environment for dynamic DNA nanotechnology. *J. R. Soc. Interface* **12,** 20150580 (2015).

118. Lakin, M. R., Youssef, S., Polo, F., Emmott, S. & Phillips, A. Visual DSD: a design and analysis tool for DNA strand displacement systems. *Bioinforma. Oxf. Engl.* **27,** 3211–3213 (2011).

119. Schaeffer, J. M. Stochastic simulation of the kinetics of multiple interacting nucleic acid strands. (California Institute of Technology, 2013).

120. Darmostuk, M., Rimpelova, S., Gbelcova, H. & Ruml, T. Current approaches in SELEX: An update to aptamer selection technology. *Biotechnol. Adv.* **33,** 1141–1161 (2015).

121. Zhou, J. & Rossi, J. Aptamers as targeted therapeutics: current potential and challenges. *Nat. Rev. Drug Discov.* **16,** 181–202 (2017).

122. Zhang, Y., Hong, H. & Cai, W. Tumor-targeted drug delivery with aptamers. *Curr. Med. Chem.* **18,** 4185–4194 (2011).

123. Tan, W. *et al.* Molecular aptamers for drug delivery. *Trends Biotechnol.* **29,** 634–640 (2011).

124. Sun, H., Tan, W. & Zu, Y. Aptamers: versatile molecular recognition probes for cancer detection. *The Analyst* **141,** 403–415 (2016).

125. Parashar, A. Aptamers in Therapeutics. *J. Clin. Diagn. Res. JCDR* **10,** BE01-06 (2016).

126. Jungmann, R. *et al.* Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat. Methods* **11,** 313–318 (2014).

127. Lin, C. *et al.* Submicrometre geometrically encoded fluorescent barcodes self-assembled from DNA. *Nat. Chem.* **4,** 832–839 (2012).

128. Modi, S. *et al.* A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol.* **4,** 325–330 (2009).

129. Sancar, A., Lindsey-Boltz, L. A., Unsal-Kaçmaz, K. & Linn, S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* **73,** 39–85 (2004).

130. Baranovskii, A. G., Buneva, V. N. & Nevinsky, G. A. Human deoxyribonucleases. *Biochem. Biokhimiia* **69,** 587–601 (2004).

131. Eder, P. S., DeVINE, R. J., Dagle, J. M. & Walder, J. A. Substrate Specificity and Kinetics of Degradation of Antisense Oligonucleotides by a 3' Exonuclease in Plasma. *Antisense Res. Dev.* **1,** 141–151 (1991).

132. Bright, S., Munro, A. J., Lawson, Y. A. & Coombs, R. R. The detection of alloantibodies to subpopulations of human lymphocytes: an adaptation of the indirect anti-immunoglobulin rosetting reaction (IARR). *J. Immunol. Methods* **24,** 175–182 (1978).

133. Dagle, J. M., Weeks, D. L. & Walder, J. A. Pathways of degradation and mechanism of action of antisense oligonucleotides in Xenopus laevis embryos. *Antisense Res. Dev.* **1,** 11–20 (1991).

134. Ortigão, J. F. *et al.* Antisense effect of oligodeoxynucleotides with inverted terminal internucleotidic linkages: a minimal modification protecting against nucleolytic degradation. *Antisense Res. Dev.* **2,** 129–146 (1992).

135. Stein, C. A., Tonkinson, J. L. & Yakubov, L. Phosphorothioate oligodeoxynucleotides--anti-sense inhibitors of gene expression? *Pharmacol. Ther.* **52,** 365–384 (1991).

136. Groves, B. *et al.* Computing in mammalian cells with nucleic acid strand exchange. *Nat. Nanotechnol.* **11,** 287–294 (2015).

137. Tay, C. Y., Yuan, L. & Leong, D. T. Nature-Inspired DNA Nanosensor for Real-Time *in Situ* Detection of mRNA in Living Cells. *ACS Nano* **9,** 5609–5617 (2015).

138. Keum, J.-W. & Bermudez, H. Enhanced resistance of DNA nanostructures to enzymatic digestion. *Chem. Commun.* 7036 (2009). doi:10.1039/b917661f

139. Mei, Q. *et al.* Stability of DNA Origami Nanoarrays in Cell Lysate. *Nano Lett.* **11,** 1477–1482 (2011).

140. Gumbiner, B. M. Cell Adhesion: The Molecular Basis of Tissue Architecture and Morphogenesis. *Cell* **84,** 345–357 (1996).

141. Stephanopoulos, N. *et al.* Bioactive DNA-peptide nanotubes enhance the differentiation of neural stem cells into neurons. *Nano Lett.* **15,** 603–609 (2015).

142. Roein-Peikar, M., Xu, Q., Wang, X. & Ha, T. Ultrasensitivity of Cell Adhesion to the Presence of Mechanically Strong Ligands. *Phys. Rev. X* **6,** (2016).

143. Chandra, R. A., Douglas, E. S., Mathies, R. A., Bertozzi, C. R. & Francis, M. B. Programmable cell adhesion encoded by DNA hybridization. *Angew. Chem. Int. Ed Engl.* **45,** 896–901 (2006).

144. Lin, D., Yurke, B. & Langrana, N. Mechanical Properties of a Reversible, DNA-Crosslinked Polyacrylamide Hydrogel. *J. Biomech. Eng.* **126,** 104 (2004).

145. Aldaye, F. A., Senapedis, W. T., Silver, P. A. & Way, J. C. A structurally tunable DNA-based extracellular matrix. *J. Am. Chem. Soc.* **132,** 14727–14729 (2010).

146. Jiang, F. X., Yurke, B., Schloss, R. S., Firestein, B. L. & Langrana, N. A. Effect of Dynamic Stiffness of the Substrates on Neurite Outgrowth by Using a DNA-Crosslinked Hydrogel. *Tissue Eng. Part A* **16,** 1873–1889 (2010).

147. Xue, F. *et al.* Development of DNA Based Active Macro–Materials for Biology and Medicine: A Review. in *Biomaterials Science and Engineering* (ed. Pignatello, R.) (InTech, 2011). doi:10.5772/23891

148. Jiang, F. X., Yurke, B., Schloss, R. S., Firestein, B. L. & Langrana, N. A. The relationship between fibroblast growth and the dynamic stiffnesses of a DNA crosslinked hydrogel. *Biomaterials* **31,** 1199–1212 (2010).

149. Rammensee, S., Kang, M. S., Georgiou, K., Kumar, S. & Schaffer, D. V. Dynamics of Mechanosensitive Neural Stem Cell Differentiation. *Stem Cells Dayt. Ohio* **35,** 497–506 (2017).

150. Lu, C.-H., Guo, W., Hu, Y., Qi, X.-J. & Willner, I. Multitriggered Shape-Memory Acrylamide-DNA Hydrogels. *J. Am. Chem. Soc.* **137,** 15723–15731 (2015).

151. Hu, Y., Guo, W., Kahn, J. S., Aleman-Garcia, M. A. & Willner, I. A Shape-Memory DNA-Based Hydrogel Exhibiting Two Internal Memories. *Angew. Chem. Int. Ed Engl.* **55,** 4210–4214 (2016).

152. Li, J. *et al.* Functional nucleic acid-based hydrogels for bioanalytical and biomedical applications. *Chem. Soc. Rev.* **45,** 1410–1431 (2016).

153. Yang, H., Liu, H., Kang, H. & Tan, W. Engineering target-responsive hydrogels based on aptamer-target interactions. *J. Am. Chem. Soc.* **130,** 6320–6321 (2008).

154. Xiong, X. *et al.* Responsive DNA-based hydrogels and their applications. *Macromol. Rapid Commun.* **34,** 1271–1283 (2013).

155. Lin, D. C., Yurke, B. & Langrana, N. A. Inducing Reversible Stiffness Changes in DNA-crosslinked Gels. *J. Mater. Res.* **20,** 1456–1464 (2005).

156. Zhou, X. *et al.* Reversibly tuning the mechanical properties of a DNA hydrogel by a DNA nanomotor. *Chem. Commun. Camb. Engl.* **52,** 10668–10671 (2016).

157. Samanta, A., Deng, Z. & Liu, Y. Infrared emitting quantum dots: DNA conjugation and DNA origami directed self-assembly. *Nanoscale* **6,** 4486 (2014).

158. Zhao, Z., Liu, Y. & Yan, H. Organizing DNA Origami Tiles into Larger Structures Using Preformed Scaffold Frames. *Nano Lett.* **11,** 2997–3002 (2011).

159. Deeds, E. J., Bachman, J. A. & Fontana, W. Optimizing ring assembly reveals the strength of weak interactions. *Proc. Natl. Acad. Sci.* **109,** 2348–2353 (2012).

160. Saccà, B., Meyer, R., Feldkamp, U., Schroeder, H. & Niemeyer, C. M. High-Throughput, Real-Time Monitoring of the Self-Assembly of DNA Nanostructures by FRET Spectroscopy. *Angew. Chem. Int. Ed.* **47,** 2135–2137 (2008).

161. Jiang, S., Yan, H. & Liu, Y. Kinetics of DNA Tile Dimerization. *ACS Nano* **8,** 5826–5832 (2014).

162. Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *AJP Cell Physiol.* **300,** C723–C742 (2011).

163. Wales, D. J. & Bogdan, T. V. Potential Energy and Free Energy Landscapes. *J. Phys. Chem. B* **110,** 20765–20776 (2006).

164. Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins Struct. Funct. Genet.* **21,** 167–195 (1995).

165. Pandey, S. *et al.* Algorithmic design of self-folding polyhedra. *Proc. Natl. Acad. Sci.* **108,** 19885–19890 (2011).

166. Gröschel, A. H. *et al.* Precise hierarchical self-assembly of multicompartment micelles. *Nat. Commun.* **3,** 710 (2012).

167. Delic, M., Göngrich, R., Mattanovich, D. & Gasser, B. Engineering of Protein Folding and Secretion—Strategies to Overcome Bottlenecks for Efficient Production of Recombinant Proteins. *Antioxid. Redox Signal.* **21,** 414–437 (2014).

168. Zimmermann, V., Beller, M. & Kragl, U. Modelling the Reaction Course of a Dynamic Kinetic Resolution of Amino Acid Derivatives: Identifying and Overcoming Bottlenecks. *Org. Process Res. Dev.* **10,** 622–627 (2006).

169. Mohammed, A. M. & Schulman, R. Directing Self-Assembly of DNA Nanotubes Using Programmable Seeds. *Nano Lett.* **13,** 4006–4013 (2013).

170. Lai, Y.-T., King, N. P. & Yeates, T. O. Principles for designing ordered protein assemblies. *Trends Cell Biol.* **22,** 653–661 (2012).

171. Chen, H. *et al.* Understanding the Mechanical Properties of DNA Origami Tiles and Controlling the Kinetics of Their Folding and Unfolding Reconfiguration. *J. Am. Chem. Soc.* **136,** 6995–7005 (2014).

172. Li, Z. *et al.* Molecular Behavior of DNA Origami in Higher-Order Self-Assembly. *J. Am. Chem. Soc.* **132,** 13545–13552 (2010).

173. SantaLucia, J. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl. Acad. Sci.* **95,** 1460–1465 (1998).

174. Nangreave, J., Yan, H. & Liu, Y. DNA Nanostructures as Models for Evaluating the Role of Enthalpy and Entropy in Polyvalent Binding. *J. Am. Chem. Soc.* **133,** 4490–4497 (2011).

175. Yoo, J. & Aksimentiev, A. In situ structure and dynamics of DNA origami determined through molecular dynamics simulations. *Proc. Natl. Acad. Sci.* **110,** 20099–20104 (2013).

176. Bai, X. -c., Martin, T. G., Scheres, S. H. W. & Dietz, H. Cryo-EM structure of a 3D DNA-origami object. *Proc. Natl. Acad. Sci.* **109,** 20012–20017 (2012).

177. Tamhane, A. G. & Dunlop, D. D. Inferences for Proportions and Count Data. in *Statistics and Data Analysis from Elementary to Intermediate* 300–301 (Prentice-Hall, 2000).

178. Legewie, S., Dienst, D., Wilde, A., Herzel, H. & Axmann, I. M. Small RNAs Establish Delays and Temporal Thresholds in Gene Expression. *Biophys. J.* **95,** 3232–3238 (2008).

179. Zaslaver, A. *et al.* Just-in-time transcription program in metabolic pathways. *Nat. Genet.* **36,** 486–491 (2004).

180. Dirks, R. M. & Pierce, N. A. Triggered amplification by hybridization chain reaction. *Proc. Natl. Acad. Sci.* **101,** 15275–15278 (2004).

181. Yin, P., Choi, H. M. T., Calvert, C. R. & Pierce, N. A. Programming biomolecular self-assembly pathways. *Nature* **451,** 318–322 (2008).

182. Caldorera-Moore, M. & Peppas, N. A. Micro- and nanotechnologies for intelligent and responsive biomaterial-based medical systems. *Adv. Drug Deliv. Rev.* **61,** 1391–1401 (2009).

183. Santini, Jr., J. T., Richards, A. C., Scheidt, R., Cima, M. J. & Langer, R. Microchips as Controlled Drug-Delivery Devices. *Angew. Chem. Int. Ed.* **39,** 2396–2407 (2000).

184. Staples, M., Daniel, K., Cima, M. J. & Langer, R. Application of Micro- and Nano-Electromechanical Devices to Drug Delivery. *Pharm. Res.* **23,** 847–863 (2006).

185. Kapsner, K. & Simmel, F. C. Partitioning Variability of a Compartmentalized *In Vitro* Transcriptional Thresholding Circuit. *ACS Synth. Biol.* **4,** 1136–1143 (2015).

186. Takahashi, K., Yaegashi, S., Kameda, A. & Hagiya, M. Chain Reaction Systems Based on Loop Dissociation of DNA. in *DNA Computing* (eds. Carbone, A. & Pierce, N. A.) **3892,** 347–358 (Springer Berlin Heidelberg, 2006).

187. Zhang, C., Yang, J. & Xu, J. Circular DNA Logic Gates with Strand Displacement. *Langmuir* **26,** 1416–1419 (2010).

188. Zhang, D. Y. & Seelig, G. DNA-Based Fixed Gain Amplifiers and Linear Classifier Circuits. in *DNA Computing and Molecular Programming* (eds. Sakakibara, Y. & Mi, Y.) **6518,** 176–186 (Springer Berlin Heidelberg, 2011).

189. Srinivas, N. Programming chemical kinetics: engineering dynamic reaction networks with DNA strand displacement. (California Institute of Technology, 2015).

190. Soloveichik, D., Seelig, G. & Winfree, E. DNA as a universal substrate for chemical kinetics. *Proc. Natl. Acad. Sci.* **107,** 5393–5398 (2010).

191. Chen, Y.-J. *et al.* Programmable chemical controllers made from DNA. *Nat. Nanotechnol.* **8,** 755–762 (2013).

192. Qian, L. & Winfree, E. Parallel and Scalable Computation and Spatial Dynamics with DNA-Based Chemical Reaction Networks on a Surface. in *DNA Computing and Molecular Programming* (eds. Murata, S. & Kobayashi, S.) **8727,** 114–131 (Springer International Publishing, 2014).

193. Wang, J. S. & Zhang, D. Y. Simulation-guided DNA probe design for consistently ultraspecific hybridization. *Nat. Chem.* **7,** 545–553 (2015).

194. Zhang, Z., Chen, N., Li, S., Battig, M. R. & Wang, Y. Programmable Hydrogels for Controlled Cell Catch and Release Using Hybridized Aptamers and Complementary Sequences. *J. Am. Chem. Soc.* **134,** 15716–15719 (2012).

195. Gatto, B., Palumbo, M. & Sissi, C. Nucleic Acid Aptamers Based on the G-Quadruplex Structure: Therapeutic and Diagnostic Potential. *Curr. Med. Chem.* **16,** 1248–1265 (2009).

196. Huizenga, D. E. & Szostak, J. W. A DNA Aptamer That Binds Adenosine and ATP. *Biochemistry (Mosc.)* **34,** 656–665 (1995).

197. Puglisi, J. D. & Tinoco, I. [22] absorbance melting curves of RNA. in *Methods in Enzymology* **180,** 304–325 (Elsevier, 1989).

198. Battig, M. R., Soontornworajit, B. & Wang, Y. Programmable Release of Multiple Protein Drugs from Aptamer-Functionalized Hydrogels via Nucleic Acid Hybridization. *J. Am. Chem. Soc.* **134,** 12410–12413 (2012).

199. Rudchenko, M. *et al.* Autonomous molecular cascades for evaluation of cell surfaces. *Nat. Nanotechnol.* **8,** 580–586 (2013).

200. He, X., Wei, B. & Mi, Y. Aptamer based reversible DNA induced hydrogel system for molecular recognition and separation. *Chem. Commun.* **46,** 6308 (2010).

201. Li, F., Zhang, H., Lai, C., Li, X.-F. & Le, X. C. A Molecular Translator that Acts by Binding-Induced DNA Strand Displacement for a Homogeneous Protein Assay. *Angew. Chem. Int. Ed.* **51,** 9317–9320 (2012).

202. Cherepanova, A. *et al.* Immunochemical assay for deoxyribonuclease activity in body fluids. *J. Immunol. Methods* **325,** 96–103 (2007).

203. Kocabey, S. *et al.* Cellular Uptake of Tile-Assembled DNA Nanotubes. *Nanomaterials* **5,** 47–60 (2014).

204. Graugnard, E. *et al.* Operation of a DNA-Based Autocatalytic Network in Serum. in *DNA Computing and Molecular Programming* (eds. Sakakibara, Y. & Mi, Y.) **6518,** 83–88 (Springer Berlin Heidelberg, 2011).

205. Barra, G. B. *et al.* EDTA-mediated inhibition of DNases protects circulating cell-free DNA from ex vivo degradation in blood samples. *Clin. Biochem.* **48,** 976–981 (2015).

206. Goltry, S. *et al.* DNA topology influences molecular machine lifetime in human serum. *Nanoscale* **7,** 10382–10390 (2015).

207. Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat. Biotechnol.* **28,** 1208–1212 (2010).

208. Hemphill, J. & Deiters, A. DNA Computation in Mammalian Cells: MicroRNA Logic Operations. *J. Am. Chem. Soc.* **135,** 10512–10518 (2013).

209. Lazarides, E. & Lindberg, U. Actin Is the Naturally Occurring Inhibitor of Deoxyribonuclease I. *Proc. Natl. Acad. Sci.* **71,** 4742–4746 (1974).

210. Mannherz, H. G., Goody, R. S., Konrad, M. & Nowak, E. The Interaction of Bovine Pancreatic Deoxyribonuclease I and Skeletal Muscle Actin. *Eur. J. Biochem.* **104,** 367–379 (1980).

211. Hahn, J., Wickham, S. F. J., Shih, W. M. & Perrault, S. D. Addressing the Instability of DNA Nanostructures in Tissue Culture. *ACS Nano* **8,** 8765–8775 (2014).

212. Zagorovsky, K., Chou, L. Y. T. & Chan, W. C. W. Controlling DNA-nanoparticle serum interactions. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 13600–13605 (2016).

213. Shaw, J.-P., Kent, K., Bird, J., Fishback, J. & Froehler, B. Modified deoxyoligonucleotides stable to exonuclease degradation in serum. *Nucleic Acids Res.* **19,** 747–750 (1991).

214. Uhlmann, E., Ryte, A. & Peyman, A. Studies on the Mechanism of Stabilization of Partially Phosphorothioated Oligonucleotides Against Nucleolytic Degradation. *Antisense Nucleic Acid Drug Dev.* **7,** 345–350 (1997).

215. Rebowski, G. *et al.* Antisense hairpin loop oligonucleotides as inhibitors of expression of multidrug resistance-associated protein 1: their stability in fetal calf serum and human plasma. *Acta Biochim. Pol.* **48,** 1061–1076 (2001).

216. Abdelgany, A., Wood, M. & Beeson, D. Hairpin DNAzymes: a new tool for efficient cellular gene silencing. *J. Gene Med.* **9,** 727–738 (2007).

217. Matsuno, H., Furusawa, H. & Okahata, Y. Kinetic studies of DNA cleavage reactions catalyzed by an ATP-dependent deoxyribonuclease on a 27-MHz quartz-crystal microbalance. *Biochemistry (Mosc.)* **44,** 2262–2270 (2005).

218. Owczarzy, R., Moreira, B. G., You, Y., Behlke, M. A. & Walder, J. A. Predicting stability of DNA duplexes in solutions containing magnesium and monovalent cations. *Biochemistry (Mosc.)* **47,** 5336–5353 (2008).

219. SantaLucia, J. & Hicks, D. The thermodynamics of DNA structural motifs. *Annu. Rev. Biophys. Biomol. Struct.* **33,** 415–440 (2004).

220. Keenan, T. M. & Folch, A. Biomolecular gradients in cell culture systems. *Lab. Chip* **8,** 34–57 (2008).

221. Barkai, N., Alon, U. & Leibler, S. Robust amplification in adaptive signal transduction networks. *Comptes Rendus Académie Sci. - Ser. IV - Phys.-Astrophys.* **2,** 871–877 (2001).

222. Schmitz, M. L., Weber, A., Roxlau, T., Gaestel, M. & Kracht, M. Signal integration, crosstalk mechanisms and networks in the function of inflammatory cytokines. *Biochim. Biophys. Acta* **1813,** 2165–2175 (2011).

223. Kiel, C., Yus, E. & Serrano, L. Engineering Signal Transduction Pathways. *Cell* **140,** 33–47 (2010).

224. Kim, J., Hanna, J. A., Byun, M., Santangelo, C. D. & Hayward, R. C. Designing responsive buckled surfaces by halftone gel lithography. *Science* **335,** 1201–1205 (2012).

225. Silverberg, J. L. *et al.* Origami structures with a critical transition to bistability arising from hidden degrees of freedom. *Nat. Mater.* **14,** 389–393 (2015).

226. Kaehr, B. & Shear, J. B. Multiphoton fabrication of chemically responsive protein hydrogels for microactuation. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 8850–8854 (2008).

227. Wehner, M. *et al.* An integrated design and fabrication strategy for entirely soft, autonomous robots. *Nature* **536,** 451–455 (2016).

228. Kurisawa, M., Matsuo, Y. & Yui, N. Modulated degradation of hydrogels with thermo-responsive network in relation to their swelling behavior. *Macromol. Chem. Phys.* **199,** 705–709 (1998).

229. Athas, J. C. *et al.* Enzyme-Triggered Folding of Hydrogels: Toward a Mimic of the Venus Flytrap. *ACS Appl. Mater. Interfaces* **8,** 19066–19074 (2016).

230. Postma, S. G. J., Vialshin, I. N., Gerritsen, C. Y., Bao, M. & Huck, W. T. S. Preprogramming Complex Hydrogel Responses using Enzymatic Reaction Networks. *Angew. Chem. Int. Ed Engl.* **56,** 1794–1798 (2017).

231. Miyata, T., Asami, N. & Uragami, T. A reversibly antigen-responsive hydrogel. *Nature* **399,** 766–769 (1999).

232. Ikeda, M. *et al.* Installing logic-gate responses to a variety of biological substances in supramolecular hydrogel-enzyme hybrids. *Nat. Chem.* **6,** 511–518 (2014).

233. Komatsu, H. *et al.* Supramolecular hydrogel exhibiting four basic logic gate functions to fine-tune substance release. *J. Am. Chem. Soc.* **131,** 5580–5585 (2009).

234. Tokarev, I. *et al.* Stimuli-responsive hydrogel membranes coupled with biocatalytic processes. *ACS Appl. Mater. Interfaces* **1,** 532–536 (2009).

235. Bocharova, V. *et al.* A biochemical logic approach to biomarker-activated drug release. *J. Mater. Chem.* **22,** 19709 (2012).

236. Shigemitsu, H. *et al.* Preparation of supramolecular hydrogel-enzyme hybrids exhibiting biomolecule-responsive gel degradation. *Nat. Protoc.* **11,** 1744–1756 (2016).

237. Zadorin, A. S. *et al.* Synthesis and materialization of a reaction–diffusion French flag pattern. *Nat. Chem.* (2017). doi:10.1038/nchem.2770

238. Venkataraman, S., Dirks, R. M., Rothemund, P. W. K., Winfree, E. & Pierce, N. A. An autonomous polymerization motor powered by DNA hybridization. *Nat. Nanotechnol.* **2,** 490–494 (2007).

239. Cangialosi, A. *et al.* Programmable DNA sequence-driven shape change of photopatterned hydrogels via high-degree swelling. *Science* **in revision.,** (2017).

240. Kang, H. *et al.* Near-infrared light-responsive core-shell nanogels for targeted drug delivery. *ACS Nano* **5,** 5094–5099 (2011).

241. Hu, Y. *et al.* Reversible Modulation of DNA-Based Hydrogel Shapes by Internal Stress Interactions. *J. Am. Chem. Soc.* **138,** 16112–16119 (2016).

242. Lin, H. *et al.* DNAzyme crosslinked hydrogel: a new platform for visual detection of metal ions. *Chem. Commun. Camb. Engl.* **47,** 9312–9314 (2011).

243. Yin, B.-C., Ye, B.-C., Wang, H., Zhu, Z. & Tan, W. Colorimetric logic gates based on aptamer-crosslinked hydrogels. *Chem. Commun. Camb. Engl.* **48,** 1248–1250 (2012).

244. Zhu, Z. *et al.* An aptamer cross-linked hydrogel as a colorimetric platform for visual detection. *Angew. Chem. Int. Ed Engl.* **49,** 1052–1056 (2010).

245. Zhu, Z. *et al.* Au@Pt nanoparticle encapsulated target-responsive hydrogel with volumetric bar-chart chip readout for quantitative point-of-care testing. *Angew. Chem. Int. Ed Engl.* **53,** 12503–12507 (2014).

246. Zhang, Z., Birkedal, V. & Gothelf, K. V. Enzyme-free colorimetric detection systems based on the DNA strand displacement competition reaction. *New J. Phys.* **18,** 055002 (2016).

247. Nutiu, R. & Li, Y. Structure-switching signaling aptamers: transducing molecular recognition into fluorescence signaling. *Chem. - Eur. J.* **10,** 1868–1876 (2004).

248. Yurke, B. & Mills, A. P. Using DNA to Power Nanostructures. *Genet. Program. Evolvable Mach.* **4,** 111–122 (2003).

249. Turberfield, A. J. *et al.* DNA Fuel for Free-Running Nanomachines. *Phys. Rev. Lett.* **90,** 118102 (2003).

250. Olson, X. *et al.* Availability: A Metric for Nucleic Acid Strand Displacement Systems. *ACS Synth. Biol.* **6,** 84–93 (2017).

251. Gawel, K. & Stokke, B. T. Logic swelling response of DNA–polymer hybrid hydrogel. *Soft Matterials* **7,** 4615 (2011).

252. Wang, X. & Ha, T. Defining Single Molecular Forces Required to Activate Integrin and Notch Signaling. *Science* **340,** 991–994 (2013).

253. Oyen, M. L. Mechanical characterisation of hydrogel materials. *Int. Mater. Rev.* **59,** 44–59 (2014).

254. Nguyen, T. D. Modeling Shape-Memory Behavior of Polymers. *Polym. Rev.* **53,** 130–152 (2013).

CURRICULUM VITAE

Joshua Fern was born in Hudson, Wisconsin on July 30th, 1988 to Donna and Gregory Fern. In May 2007, he graduated salutatorian from Hudson Senior High School. From 2007 to 2012, Joshua studied Chemical and Biological Engineering at the University of Wisconsin-Madison. He received B.S. degrees in Chemical and Biological Engineering and in Chemistry, along with a Certificate of Biology in Engineering, in August 2012. While at UW-Madison, Joshua spent two years researching the effect of the cell's innate immune response on virus infection spread across cell monolayers in the laboratory of Professor John Yin. In the summer after his fourth year of B.S. study, Joshua worked as an intern at BASF, SE in Ludwigshafen, Germany where he investigated the effect of reactor design and reactor parameters on the size distributions of photopolymerized particles. In the fall of 2012, Joshua began his graduate studies at Johns Hopkins University in the Department of Chemical and Biomolecular Engineering and started his research in the laboratory of Professor Rebecca Schulman. At Johns Hopkins, Joshua also served as a teaching assistant for the "Modeling, Dynamics and Control of Chemical and Biological Processes" course.

Technical Skills

MOLECULAR BIOLOGY

·DNA purification methods ·Agarose and polyacrylamide gel electrophoresis ·Mammalian/bacterial cell culture ·Cell transfection ·Virus infection assays

BIOMATERIALS ENGINEERING

·Design and characterization of DNA nanostructures, circuits, and sensors ·Synthesis of DNA-peptide and DNA-surface conjugates ·Surface modification chemistry ·Single-molecule assay development ·Photolithography

IMAGING/QUANTITATIVE TECHNIQUES

·Live cell imaging ·Fluorescence microscopy ·Atomic force microscopy (AFM) ·Absorbance spectroscopy ·Fluorescence spectroscopy

PROGRAMMING/DATA ANALYSIS

·Expert data and image analysis using MATLAB ·Deterministic/stochastic chemical reaction simulations ·Reaction parameter fitting ·ImageJ ·Inkscape ·Intermediate Javascript/Python ·C++/C# and HTML/CSS

Work Experience

Johns Hopkins University, Baltimore, MD 2012 - present
Research Assistant *Advisor: Dr. Rebecca Schulman*

- Designed fluorescence and atomic force microscopy assays to quantify DNA origami nanostructure self-assembly and interactions with mammalian cells.
- Developed novel DNA-based molecular circuits for time-release applications and increased lifetimes in physiological environments.
- Developed numerical algorithms and models in MATLAB to simulate DNA circuit reactions, fit reaction rate constants, and model the effect of nucleases on circuit performance.
- Developed novel mechanisms to connect DNA-based circuits with hydrogels for chemo-mechanical sensors and morphogenic materials.
- Mentored and trained multiple students in efficient project management, experiment design, and data analysis.
- Organized a multi-week lab renovation without interim loss of research capabilities.
- Managed laboratory chemical and biological safety compliance.

BASF SE, Ludwigshafen, Germany June 2011 - Sept. 2011
Polymer Process Development Engineer

- Developed pre-pilot plant reactor designs for the photopolymerization of polymer beads with post-production purification.
- Compared methods of determining reactor-dependent polymer bead size distribution.
- Communicated project details and goals to coworkers with varying levels of English-proficiency using both German and English.

University of Wisconsin-Madison, Madison, WI 2009 - 2011
Research Assistant *Advisor: Dr. John Yin*

- Analyzed virus infection spread across cell monolayers using fluorescence microscopy.
- Generated fluorophore-producing cell lines using transfection and cell sorting techniques.
- Developed assays to quantify the effect of interferon on the spread of virus infection.

Education

Doctor of Philosophy in Chemical and Biomolecular Engineering
May 2017
Advisor: Prof. Rebecca Schulman
The Johns Hopkins University, *Baltimore, MD*
Thesis: Design and applications of DNA-based devices for self-assembly, molecular circuits, and soft materials.
Areas: DNA, nanotechnology, hydrogels, molecular circuits, self-assembly

Bachelor of Science Double Major: Chemical Engineering and Chemistry
August 2012
University of Wisconsin-Madison, *Madison, WI* GPA: 3.75/4.0
Certificate of Biology in Engineering

Publications

"Modular DNA strand-displacement controllers that direct material shape and property changes." Fern, J. and Schulman, R., **in preparation**.

"Design and characterization of DNA strand-displacement reactions in serum-supplemented cell medium." Fern, J. and Schulman, R., *ACS Synthetic Biology*. **2017**. DOI: 10.1021/acssynbio.7b00105.

"Stable DNA-based reaction-diffusion patterns." Zenk, J.; Scalise, D.; Wang, K.; Dorsey, P.; Fern, J.; Cruz, A.; and Schulman, R., *RSC Advances*. **2017**, 7, 18032-18040.

"DNA strand-displacement timer circuits." Fern, J.; Scalise, D.; Cangialosi, A.; Howie, D.; Potters, L.; Schulman, R., *ACS Synthetic Biology*. **2016**, 6, 190-193.

"The energy landscape for the self-assembly of a two-dimensional DNA origami complex." Fern, J.; Lu, J.; Schulman, R., *ACS Nano*. **2016**, 10(2), 1836-1844.

"Electrostatic-driven lamination and untwisting of β-sheet assemblies." Hu, Y.; Lin, R.; Zhang, P.; Fern, J.; Cheetham, A.; Patel, K.; Schulman, R.; Kan, C.; Cui, H., *ACS Nano*. **2016**, 10(1), 880-888.

Awards and Honors

- Graduate Student & Postdoctoral Fellow Research & Education Award, Johns Hopkins University, 2017.
- Carole Foster Scholarship, University of Wisconsin-Madison, 2011.
- Dean's List, University of Wisconsin-Madison, 2007-2012.
- Academic Excellence Scholar, University of Wisconsin-Madison, 2007-2012.

Research Mentoring Experience

Chad Fischer Summer 2016
Research Topic: Developing DNA-based circuits to sense ATP or thrombin
Current Position: Undergraduate at Rice University

Jen Lu 2014-2015
Research Topic: Measuring the temperature-dependent self-assembly of DNA origami nanostructures
Current Position: Ph.D. candidate in biomedical engineering at Johns Hopkins University